Received December 8, 2016; accepted December 27, 2016, date of publication March 21, 2017, date of current version April 24, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2680318

Planning Wireless Cellular Networks of Future: Outlook, Challenges and Opportunities

AZAR TAUFIQUE[1], MONA JABER[2], ALI IMRAN[3], ZAHER DAWY[4], AND ELIAS YACOUB[5]

[1]TechTrained LLC, Richardson, TX 75080, USA
[2]5G Innovations Center, University of Surrey, Guildford GU2 7XH, U.K.
[3]Department of Electrical and Computer Engineering, University of Oklahoma, Tulsa, OK, 74135 USA
[4]Department of Electrical and Computer Engineering, American University of Beirut, Beirut 11072020, Lebanon
[5]Arab Open University, Beirut 20584518, Lebanon

Corresponding author: A. Taufique (azar.taufique@techtrained.com)

This work was supported by the Computer and Network Systems Program of the National Science Foundation under Grant 1559483.

ABSTRACT Cell planning (CP) is the most important phase in the life cycle of a cellular system as it determines the operational expenditure, capital expenditure, as well as the long-term performance of the system. Therefore, it is not surprising that CP problems have been studied extensively for the past three decades for all four generations of cellular systems. However, the fact that small cells, a major component of future networks, are anticipated to be deployed in an impromptu fashion makes CP for future networks *vis-a-vis* 5G a conundrum. Furthermore, in emerging cellular systems that incorporate a variety of different cell sizes and types, heterogeneous networks (HetNets), energy efficiency, self-organizing network features, control and data plane split architectures (CDSA), massive multiple input multiple out (MIMO), coordinated multipoint (CoMP), cloud radio access network, and millimetre-wave-based cells plus the need to support Internet of Things (IoT) and device-to-device (D2D) communication require a major paradigm shift in the way cellular networks have been planned in the past. The objective of this paper is to characterize this paradigm shift by concisely reviewing past developments, analyzing the state-of-the-art challenges, and identifying future trends, challenges, and opportunities in CP in the wake of 5G. More specifically, in this paper, we investigate the problem of planning future cellular networks in detail. To this end, we first provide a brief tutorial on the CP process to identify the peculiarities that make CP one of the most challenging problems in wireless communications. This tutorial is followed by a concise recap of past research in CP. We then review key findings from recent studies that have attempted to address the aforementioned challenges in planning emerging networks. Finally, we discuss the range of technical factors that need to be taken into account while planning future networks and the promising research directions that necessitates the paradigm shift to do so.

INDEX TERMS HetNets planning, energy efficient planning, 5G network planning.

I. INTRODUCTION

Research in cellular planning (CP) is older than the cellular system itself [1]–[3]. However, the first generation of cellular systems were planned almost manually as the focus was on providing coverage to serve the elite of society only. The gigantic subscription fees, low traffic loads, lack of competition and relative abundance of spectrum at that time meant not much effort had to be invested to optimize the network plan. As the trend moved toward ubiquity of cellular service, the foremost optimization objective that emerged was to maximize the coverage while keeping the number of base stations at a minimum [4]–[6]. This prompted the first call for CP optimization techniques to be investigated and raised the need for automated computer-aided CP tools [7], [8] thereby triggering the academic and industrial research in this area that has grown continuously thereafter. We can refer to this initial era of CP research as classic CP that roughly spanned over the decade of the '90s. Research in this *classic CP* era can be broadly described as being mainly focused on optimizing the location and number of base stations (BS) while largely abstracting the parameter optimization of the base station themselves.

Introduction of data services and, consequently, crowded networks at the beginning of the new millennium meant that operators had to tweak and optimize a large number of BS parameters in the planning process to squeeze out all possible bits of capacity [9], [10]. This strategy shifted the focus of CP research from classic and relatively primitive to a more advanced planning [11] approach that we can refer to as holistic CP. In holistic CP, in addition to BS locations, BS parameters such as number of sectors, azimuths, tilts, transmission powers, pilot powers etc. were considered while formulating and optimizing the CP problem [12]. This era of holistic CP can be roughly mapped to the first two to three quarters of the last decade. Gould [13] described some challenges that CP engineers faced when doing holistic CP. While holistic CP solutions with reasonable computation complexity were still being sought [14]–[16], the advent of LTE and LTE-advanced (LTE-A) at the beginning of current decade again called for a major revamp of the CP paradigm. Unprecedented demand for higher data rates combined with projected proliferation of internet of things (IoT) mean new technologies such as massive MIMO, smart femto cells [17], [18], fractional frequency reuse [19], CoMP, C-RAN, and mmWave had to be resorted to in emerging networks. While adaptation of these technologies in 5G offers promising avenues for raising cellular system capacity, it put forth a whole new set of challenges to the CP research community. In addition, in the wake of the rising cost of energy and environmental awareness, energy efficiency became a newly added constraint to the CP problem that asks for a significantly different, if not totally new, approach toward CP [20]. Furthermore, in the emerging socio-economic structure, the average revenue per bit earned by the operators is diminishing. This trend is pushing operators to rely on self organization network (SON) features to minimize OPEX and CAPEX. While SON is a promising paradigm that improves capacity, and reduces total cost of ownership (TCO) for the network operators [21]. However, it remains unclear how the two paradigms CP and SON, CP being too old and SON being too young, will fit together in emerging networks such as 5G.

Regardless of physical layer waveform and spectrum band adapted for 5G cellular systems, it is clear that the majority of the 1000x target capacity gain must come from network densification. Realizing the massive potential of network densification by small cells, industry pundits have been forecasting an explosive growth of small cells for the past few years. However, to date, mass deployments of small cells remain elusive mainly due to the fact that the ultra-dense deployment of small cells comes with its own set of peculiar planning challenges. The key challenge being *how to plan and roll out a heterogenous network that will contain unplanned deployment of small cells.*

In the backdrop of these recent developments, this paper aims to analyse the state of the art in CP and identify the challenges and opportunities therein in context of emerging cellular networks such as 5G. Though this article provides a concise review of selected literature on CP, its main objective is not to provide comprehensive survey of literature. Instead this article has following goals:

1) Provide a brief tutorial on CP process to highlight the conflicting objectives and constraints that make CP one of the most challenging problems in wireless communications. (Section II)
2) Provide concise and tabular recap of literature to guide the reader to sources that have addressed different parts of CP to date. (Section III)
3) Identify recent trends in CP such as energy-focused planning, planning for traffic uncertainties and with CoMP in mind. (Section IV)
4) Provide an overview of the models and techniques that have emerged recently to improve CP such as models for cell load, interference, BS location randomness, channel variation, and TCO. (Section V)
5) Identify the technical factors that make planning a heterogenous network (HetNet) different from macro-cell-only networks and how these factors can be accounted for in a new HetNet planning paradigm. (Section VI)
6) Finally, identify the prospects, challenges, and opportunities that lie in the CP paradigm in wake of 5G and beyond. (Section VII). This section answers the questions: what, why and how the CP paradigm will change with the advent of C-RAN, M2M, D2D, and CDSA, Massive MIMO, and mmWave in 5G and beyond.

   Fig. 1 shows the contents, contribution, and layout of this paper's details

II. A BRIEF TUTORIAL ON PLANNING PROCESS

The cell planning process consists of three phases: pre-planning, or dimensioning; detailed planning; and post planning, or optimization, as shown in Fig. 2. The output of the dimensioning phase is an approximate number of BSs required to cover an area of interest. For a recent tutorial on the dimensioning process, readers can refer to [22]. The detailed planning phase allows determining the actual positions of the BSs within the area to be served. In the optimization phase, which occurs after the network has been deployed and is running, the network performance is analyzed, potential problems detected, and improvements made to enhance network operation [23]. This paper focuses on the detailed planning phase. In this section, we briefly review the CP process, including its objectives, input and output parameters to the optimization process and to the CP phases. We also give a simple analysis of the complexity of the CP problem. This section aims to provide the necessary background needed for the discussion in successive sections.

A. CELL PLANNING OBJECTIVES

The objectives of CP heavily depends on the business strategy of the operators. The coverage target for different services, the pricing and throughput policies, regulatory constraints, market share goals and competition are some factors among many that define the CP objectives. Ultimately, CP objectives can be boiled down to the following set of optimization targets identified in the cell planning problem:

1) **Minimize TCO.** In addition to minimizing the overall network cost, this objective may also include

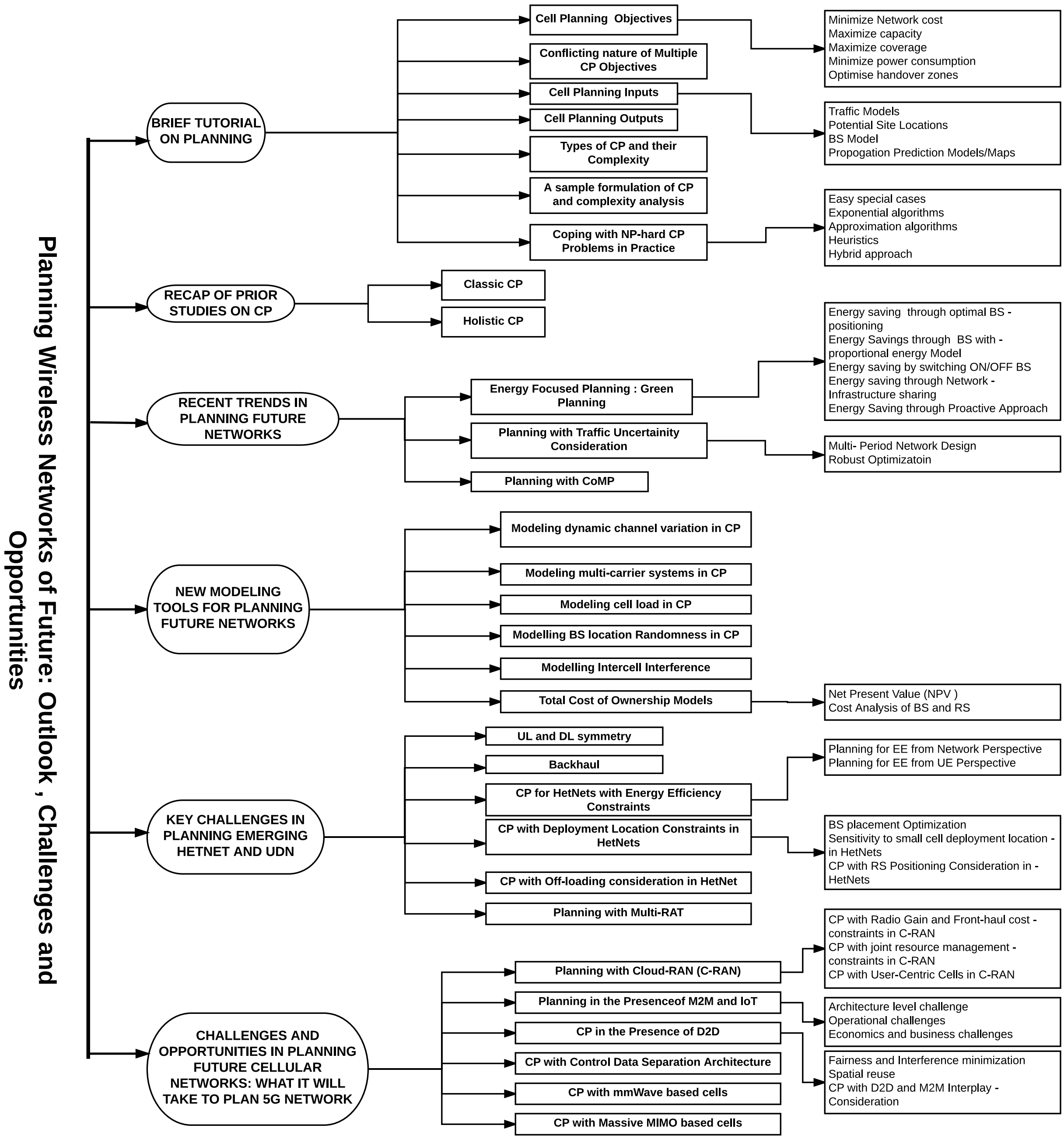

FIGURE 1. Layout of the contents and paper contribution: The flow chart explains challenges, opportunities, dependencies and factors in planning future cellular networks.

FIGURE 2. Three phases of cellular network planning and optimization.

minimizing economic costs related to deployment costs and parameter optimization.

2) **Maximize capacity.** For a single service, this objective can be defined as the number of users who can be served at one time. In the case of multi-service traffic, capacity can be approximated in terms of global throughput.
3) **Maximize coverage.** This includes satisfying coverage policy requirements for various services. Up Link (UL) and Down Link (DL) coverage must be balanced. Both traffic channels and coverage of common channels must be considered.
4) **Minimize Power Consumption.** Health concerns have motivated the radiated power minimization objective. However, recent awakening of a desire for greener

wireless systems has added more depth to this objective. Consequently, power consumption, including fixed circuit power as well as variable transmission power, must be minimized.

5) **Optimise handover (HO) zones.** In a well-planned cellular system, a certain proportion of the area of each cell should overlap with neighboring cells to satisfy HO conditions. HO zones are essential to guarantee continuity of service between the sectors. It also strengthens the radio link against fast fading and shadowing. However, too much overlap may result in wastage of power, and radio resources, and increase in interference and electro-smog, making it a tricky planning objective.

B. CONFLICTING NATURE OF MULTIPLE CP OBJECTIVES

Ironically, the CP objectives listed in Section II-A mutually conflict, hence, giving rise to the immense research on CP in past two decades. For example, maximizing the coverage and capacity requires deploying more base stations, which in turn, increases the network cost. Similarly, coverage maximization contradicts the objective of reducing power consumption and electro-smog. Regardless of which technique is adopted to solve the CP problem, competing multiple objectives need to be addressed, although this is often done implicitly rather than explicitly. The main implication of having multiple objectives is that a set of optimal solutions, rather than a single solution, is obtained. Although several alternatives to cope with multi-objectivity have been proposed in the literature, no approach proves more prevalent [24].

To cope with more than one CP objective, multi-objective functions are often defined. Previous CP studies used two different ways to represent a multi-objective function. One way is to use a linear combination of different objective criteria to form a single objective function, where different objectives are given a certain weight between 0 and 1 [25]–[29].

In the second method, the problem is formulated by a set of decision variables i.e., parameter space vector and a set of objective vectors. When there is no solution that can improve one objective without degrading the other objective, it is considered as the optimal solution. This solution is referred to the Pareto optimal solution [29]–[31]. The objective functions of interest can also be assigned weights to reflect their importance relative to each other. Such weighted multi-objective functions give more flexibility to the network planner by assigning higher or lower weight to put more or less emphasis on a given objective.

C. CELL PLANNING INPUTS

Different inputs are required to solve the cell planning difficulty depending on objectives in focus and phase of planning. Usually, the following inputs need to be known [32]:

1) **Traffic Models:** User traffic distribution is a main factor that ultimately determines the cellular system plan and, hence, is a key input in the CP process. In GSM (mono-service systems), for instance, geographical characterisation of traffic distribution is sufficient. However, with multi-service systems supporting data, traffic characterisation based on types and level of service is needed [33]. Test point based traffic models are often used for CP traffic modeling, for the sake of practicality [6], [34]–[36]. In this model, an area is characterized over a time interval and all located mobile terminals are bundled into a single test point. This point represents the cumulative traffic, or traffic intensity, from all these terminals, over the determined interval.
2) **Potential Site Locations:** Theoretically, a base station can be installed anywhere. However in the real world, a set of candidate sites is first pre-determined and used as input to the CP, to incorporate the real estate constraints. The objective, thus, is to find the optimum subset of BS locations. These potential BS locations are determined by taking into account the constraints such as, socio-economic feasibility and availability of site(s), traffic density, building heights, terrain height(s) and preexistence of a site(s) by the same or other operators.
3) **BS Model:** There are many parameters that define the BS model such as: antenna type and height, receiver sensitivity, load capacity, transmit power and capital and operational costs [37]. Moreover, heterogeneous networks necessitate modeling of new types of nodes; for instance relay stations (RS), pico-cells, femto-cells, and small cells.
4) **Propagation Prediction Models/Maps:** A key input to the planning process is the signal propagation model. The potential of this model is to incorporate reflection, differentiation, absorption, and propagation of the signal in real environment. Taking into account the natural and man-made structures, vegetation and topography of an area, highly determines the accuracy of the CP outcomes [38]. Very sophisticated planning tools rely on ac- tual measurement based propagation maps, or ray tracing based complex analysis, to predict the propagation [39]. However, obtaining complete propagation maps of a large area using these methods is a very cumbersome, time consuming, and expensive process. For this reason, different empirical models have been proposed in the literature. Such models abstract the experimental and statistical data in the form of deterministic expressions, that can easily be used in the CP. Okumura [40], Hata [41], and COST 231 [42] are a few examples of such well known propagation models used in CP to depict propagation loss in different environments and scenarios. A fine tuning of these models is done by setting parameters within these models to reflect the real-world conditions as closely as possible. While propagation models for sub 5 GHz frequencies are well established, research on developing such models for higher frequencies such as mmWaves is still in progress [43].

D. CELL PLANNING OUTPUTS

The goal of the CP process is to provide one or more of the following outputs:

1) The optimal number of base stations;
2) The best locations to install base stations;
3) The types of base station optimal for each location;
4) The configuration of parameters such as antenna height, number of sectors and sector orientation, tilt, power;
5) Frequency reuse pattern;
6) Capacity dimensioning, e.g. number of carriers or carrier components per sector.

E. TYPES OF CP AND THEIR COMPLEXITY

The objectives, input and output of the CP process also depend on the type of planning. There are generally two types of CP, roll out and incremental, as explained below:

1) Roll-out CP: This is the CP where no prior networks exists and a plain state approach can be used to meet all the objectives of interest. In terms of input parameters, in this phase the traffic distribution is not exactly known yet. Estimates of traffic based on geo-marketing forecasts are used for planning in this phase
2) Incremental Planning: This type of CP is generally carried out after the first roll-out planning to meet the increasing demand. Unlike the plane state approach, planning in this phase is bounded by additional constraints imposed by existing sites. However, in this phase the traffic distribution can be modeled now with much better accuracy using the measurements from existing network reports [44]. It is anticipated that 5G deployment will mostly require incremental planning by building on LTE/UMTS/GSM network.

Both CP types correspond to the second part of Fig. 2, with incremental planning also touching on the third part (post deployment optimization).

F. A SAMPLE FORMULATION OF CP AND COMPLEXITY ANALYSIS

For CP, if the main objective is to maximize service area fairness, maximize capacity and minimize power consumption in the system. The objective function can be modelled for a number of constraints. Using the notation defined in Table 1, the problem of holistically planning a cellular network can now be formulated as multi-objective optimisation problem as below:

$$
\begin{aligned}
&\max_{\mathcal{Q}_b,\mathcal{Q}_r,\mathcal{H}_r,\mathcal{H}_b,\mathcal{S}_b,\mathcal{P}_s,\mathcal{P}_r,\mathcal{F}_s,\boldsymbol{\phi}^s,\boldsymbol{\theta}} f(\Gamma,\Upsilon,\Omega)\\
&\text{subject to: } B \leqslant B_{max}\\
&\quad R \leqslant R_{max}\\
&\quad 1 \leqslant S_b \leqslant S_{b,max}, \quad \forall S_b \in \mathcal{S}_b\\
&\quad \frac{360}{S_b} i - k_{\phi,max}\delta_\phi \leqslant \phi^s \leqslant \frac{360}{S_b} i + k_{\phi,max}\delta_\phi,\\
&\qquad\qquad i = 1,2,3\ldots S_b\\
&\quad 0 \leqslant \theta^s \leqslant k_{\theta,max}\delta_\theta, \quad \forall \theta^s \in \boldsymbol{\theta}\\
&\quad h_{s,min} \leqslant h_s \leqslant h_{s,max}, \quad \forall h_s \in \mathcal{H}_s\\
&\quad h_{r,min} \leqslant h_r \leqslant h_{r,max}, \quad \forall h_r \in \mathcal{H}_r\\
&\quad p_{s,min} \leqslant p_s \leqslant p_{s,max}, \quad \forall p_s \in \mathcal{P}_s\\
&\quad p_{r,min} \leqslant p_r \leqslant p_{r,max}, \quad \forall p_r \in \mathcal{P}_r
\end{aligned}
\tag{1}
$$

The above formulation can help to gauge the unfathomableness of the solution space of the CP problem. Taking a toy

TABLE 1. Symbol description.

Symbol	Description
b	b^{th} base station
$\mathcal{B}$	set of all base stations in systems
B	total number of BS i.e. $\lvert\mathcal{B}\rvert = B$
B_{max}	maximum number of BS that can be afforded.
A	Total area of interest
$\mathcal{Q}$	set of Q bins that constitute A
q	q^{th} bin, $\sum_{i=1}^{Q} q_i = A$, & $\frac{A}{Q} = q, \forall q \in \mathcal{Q}$
$\mathcal{Q}_b$	set of bins in which BS are located, $\mathcal{Q}_b \subseteq Q$
$\mathcal{S}$	set of all sectors in the systems
S	total number of sectors in system i.e. $\lvert\mathcal{S}\rvert = S$
s	denotes s^{th} sector
S_b	total number of sectors b^{th} BS has
$\mathcal{S}_b$	$\mathcal{S}_b = \{S_1, S_2, S_3...S_B\}$, $S = \lvert\mathcal{S}\rvert = \sum_{b=1}^{B} S_b$
$S_{b,max}$	maximum number of sectors a BS can have
h_s	(antenna) height of s^{th} sector antenna on BS
$\mathcal{H}_s$	set of all sector antenna heights
$h_{s,max}$	maximum allowed h_s
$h_{s,min}$	minimum allowed h_s
δ_{h_b}	step with which h_s can vary
f_s	fractional frequency reuse factors in s^{th} sector
$\mathcal{F}$	set of f_s for all sectors.
k_f	number of different values f_s can have
$\mathcal{R}$	set of RSs in the system R i.e. $\lvert\mathcal{R}\rvert = R$
r	r^{th} RS in the system
R_{max}	maximum RSs that can be afforded
$\mathcal{Q}_r$	set of bins in which RS are located, $\mathcal{Q}_r \subseteq Q$
$\mathcal{H}_r$	set of all RS antenna heights
h_r	height of r^{th} RS antenna,
δ_{h_r}	step with which h_r can vary
$\boldsymbol{\phi}$	vector of azimuth angles of all sectors
ϕ^s	azimuth angle of s^{th} sector
$\boldsymbol{\theta}$	set of tilt angles of all sectors
θ^s	tilt angle of s^{th} sector
δ_ϕ, δ_θ	steps sizes for azimuth and tilts change
$k_{\theta,max}$	maximum steps of tilt change
$k_{\phi,max}$	maximum steps of azimuth change
$\mathcal{P}_s$	set of transmission powers of all sectors
p_s	transmission power from s^{th} sector
δ_{p_s}	step with which p_s can vary
$p_{s,max}$	maximum allowed value of p_s
$p_{s,min}$	minimum allowed value of p_s
$\mathcal{P}_r$	set of transmission powers of all RS
p_r	transmission power from r^{th} RS
δ_{p_r}	step with which p_r can vary
$p_{r,max}$	maximum allowed value of p_r
$p_{r,min}$	minimum allowed value of p_r
G_q^s	gain from the s^{th} sector antenna to q^{th} bin.
α	path loss co-efficient
β	pathloss exponent
φ_v	vertical beamwidth of the antenna
φ_h^s	horizontal beamwidth of s^{th} sector antenna
Υ	capacity wise performance indicator
Ω	total power consumption in the system
Γ	service area fairness wise key performance indicator (KPI)
$\mathcal{X}\backslash y$	means all elements of $\mathcal{X}$ except y.

example of only 56 cell cellular system and focusing on solving for optimal tilt angle only, assuming a quantization to ten possible values, a brute force based solution will have to assess $\{k_{\theta,max}+1\}^{\left(\sum_{b=1}^{B} S_b\right)} = 10^{56}$ possible solutions. With a state of the art computer having processing speed of 10^{12} evaluations per second, finding an optimal solution may require as long as $\frac{10^{56}}{10^{12}\times 8.6\times 10^{4}\times 365}$ seconds, which is clearly prohibitive. Note that the actual size of the solution space of a typical holistic planning problem represented by (1) would be even more gigantic as can be sketched by the expression below:

$$\begin{aligned} &\frac{Q!}{B!(Q-B)!}\times\frac{(Q-B)!}{R!(Q-B-R)!} \\ &\quad\times\left\{\frac{h^{s}_{b,max}-h^{s}_{b,min}}{\delta_h}+1\right\}^{\left(\sum_{b=1}^{B} S_b\right)}\times(S_{max})^{B} \\ &\quad\times\left\{\frac{h_{r,max}-h_{r,min}}{\delta_h}+1\right\}^{R} \\ &\quad\times\left\{\frac{p^{s}_{b,max}-p^{s}_{b,min}}{\delta_p}+1\right\}^{\left(\sum_{b=1}^{B} S_b\right)} \\ &\quad\times\left\{\frac{p_{r,max}-p_{r,min}}{\delta_h}+1\right\}^{R} \\ &\quad\times k_f^{\left(\sum_{b=1}^{B} S_b\right)}\times\{k_{\phi,max}\times 2+1\}^{\left(\sum_{b=1}^{B} S_b\right)} \\ &\quad\times\{k_{\theta,max}+1\}^{\left(\sum_{b=1}^{B} S_b\right)} \end{aligned} \tag{2}$$

By the preponderance of (2) it is clear that a brute force solution is theoretically not possible in reasonable computing time. Since there is no known polynomial time efficient algorithm in the literature for this or similar problem, CP problem has been shown to be NP-hard a number of times [45]–[48]. Appreciation of the complexity of cell planning problem through this analysis should help better understand rationale behind solution approaches taken in various works in the literature.

G. COPING WITH NP-HARD CP PROBLEMS IN PRACTICE

When considering the CP problems that belong to the class of NP-hard combinatorial problems, the most common approaches that have been used in literature can be classified into the following:

- **Easy special cases**. In this approach, the problem is not solved in its full generality. Rather properties of the input instances are identified and exploited that make the problem easier and mathematically tractable, and then an algorithm is designed that makes use of these properties. Although the advantages of this approach such as robustness and transparency are strongly advocated in [49] and [50], it largely remains an under explored territory in CP domain. Recent examples of use of this approach can be found in [51] and [52].
- **Somewhat efficient exponential algorithms**. Here an algorithm is designed that always solves the problem with running time not polynomial, but still much faster than exhaustive search. This approach may be useful for inputs of moderate size. Examples for use of this approach can be found in [26], [53], and [54].
- **Approximation algorithms**. In this approach the quality of solution is sacrificed to obtain more efficient algorithms. Instead of finding the optimal solution, the algorithm settles for a near optimal solution with advantage of making the problem easier. Examples of use of such approach can be found in [55]–[57].
- **Heuristics** In this approach heuristics are used to design algorithms that work well on many instances, though not on all instances. This is perhaps the approach most commonly used in practice [58] and heuristics such as simulated annealing [15], [59], genetic algorithms [47], [60]–[62], particle swarm [48], [63], Taguchi's method [64], bee colony optimisation [65], tabu-search [66] or k-mean algorithm [67] have been applied to obtain near optimal solutions for various CP problems. A detailed discussion of use of evolutionary heuristic for planning problems can be found in [61].
- **Hybrid approach** A number of hybrid approaches that combine analytical and heuristic techniques or combine more than one heuristic in cascaded stages to solve the NP-hard planning problems have also been proposed in the literature [68]. For example, authors in [69] present a hybrid approach consisting of three stages. In the first stage, a good feasible solution to the problem is found by using constraint satisfaction technique embedded with a problem-specific search guidance. The second stage is to apply a good local search procedure to improve this solution. The third stage is to make a further improvement to the solution derived from the second stage. The best objective function value obtained from the second stage is used as the upper bound, then a constraint optimization technique is applied to improve the solution. Numerical results show that optimal solutions are always obtained for small to medium sized problems. For larger sized problems, the final results are on average within 6-7 percent of the lower bounds. Such hybrid approach can be an efficient tool for tackling a wide range of combinatorial NP-hard problems.

III. RECAP OF THE PRIOR STUDIES ON CP

The objective of this article is not to provide a comprehensive survey the past of CP but characterize the future of CP by building on insights from past and present. Therefore, instead of providing a detailed review of past literature on planning, in Table II, we concisely summarize the representative research works on CP that have been carried out in time between the dawn of 3rd generation cellular system (UMTS) and the emergence of a 4th generation cellular system (LTE). In addition to the detailed classification labels given in columns of Table II, these works can broadly be classified into

- Classic CP
- Holistic CP

TABLE 2. Parameters addressed in the CP literature.

Parameters	BS Location	Frequency	Budget	Incremental	Considering Interference	Tx. power,pilot planning	Height	Heterogeneous traffic	No. of Sectors	Considering BS Antennas	Azimuth Planning	Tilt planning	2G/TDMA friendly	3G/CDMA friendly	4G/OFDMA friendly	Traffic uncertainties	energy efficiency	heterogeneous deployment	analytical models
[6], [70], [71]	✓				✓			✓					✓						
[72], [73], [71]	✓				✓		✓	✓		✓	✓	✓		✓					✓
[10], [74]					✓			✓	✓	✓	✓			✓					
[75]	✓		✓		✓	✓		✓		✓	✓	✓	✓						
[76]	✓				✓			✓		✓				✓					
[66], [77]	✓		±	✓				±					✓	±					
[8], [47]	✓							✓		✓				✓					
[60]	✓							✓					✓						
[70]	✓				✓	✓													
[49], [36]	✓				✓								✓						✓
[68], [78]	✓												✓						
[79]	✓				✓			✓					✓	✓					
[80]	✓				✓	✓								✓			✓		
[81]	✓			✓	✓	✓		✓					✓						
[82]	✓	✓			✓	✓		✓		✓	✓				✓				
[83]	✓									✓			±	±	±				
[84]	✓												±	±	±			✓	
[85]	✓							✓					±	±	±				
[16]	✓				✓									✓					
[61]	✓		✓		✓	✓	✓			✓	✓			✓					✓
[86]	✓												✓					✓	
[87]	✓				✓					✓				✓				✓	
[88], [89]	✓				✓					✓				✓					

The Classic CP was mainly concerned with optimising the number and location of base stations. As we can see in Fig. 5, in the previous technologies there were no new features added. With the widespread deployment of 2G and 3G cellular networks, explosive traffic demand, and invasive data services, the classic planning could no longer serve the objectives of CP. In order to cater to multiple objectives, such as coverage, capacity, QoS, cost of network and energy consumption, many parameters needed to be considered in the CP, which motivated the development of holistic CP.

In the Table II, we have also summarized the holistic CP parameters whose optimization have been addressed and researched in CP literature. The table provides a taxonomy of the works done in terms of BS location, frequency, budget, interference, transmission power, height, heterogeneous traffic, number of sectors, antennas, tilt planning, traffic uncertainties, and analytical models etc.

IV. RECENT TRENDS IN PLANNING FUTURE NETWORKS

In this section, we review the literature on recent trends in CP that include considerations for energy efficiency, uncertainty in traffic, and CoMP.

A. *ENERGY FOCUSED PLANNING: GREEN PLANNING*

In recent years improving energy efficiency in cellular operation has become an integral part of CP, partly to reduce carbon footprint and partly to reduce OPEX. Furthermore, as the cellular networks are becoming denser and revenues per bit are decreasing, the need for energy efficient cellular systems is growing more than ever. In the following we review several mechanisms to incorporate energy efficiency in to the CP process that have been proposed in literature.

1) ENERGY SAVINGS THROUGH OPTIMAL BS POSITIONING

Authors in [80] tackle the basic BS location problem and assignment of mobile users to appropriate BSs in 3G W-CDMA uplink environment. The authors propose a constraint satisfaction model and apply different techniques like variable ordering and value ordering to find good optimal solutions. Instead of cost minimization, the objective of their model is to minimize the total transmitted power. Once the location and power configuration of BSs are known, the next step is to study BS assignment to the higher level. This involves the investigation of an access network sub problem. It is shown that local approaches, that aim at reducing the energy consumption of individual network components, can be quite effective. However, global approaches, that consider the entire network energy consumption in the network design, planning, and management phases are a must, for a holistic approach to energy efficient networking.

Compared to previous works on energy savings via BS switching, in [90] and [91] the authors investigate the dominating factors in the energy savings. Energy consumption of the BS amounts to nearly 850W, with the energy needed to transmit from the antennas amounting only up to 40W and the rest expended even in case of idle operation. Their analysis shows that the mean and variance of traffic profile and the BS density are the dominant factors that determine the amount of achievable energy saving. Moreover, an expression is obtained that indicates that the energy saving increases when the traffic, mean/variance ratio and the number of neighboring BS have higher value. It means, for instance, that the greatest energy savings are likely to be realized in urban commercial areas (since such an area is likely to show both high traffic variance between day time and night time as well as high BS density). It is also emphasized that, the slope of traffic variation is more important than the maximum value in estimating the traffic profile because the slope directly determines the switching-on/off time.

In [92] authors investigate techniques to optimize the number of base stations and their locations for energy efficiency. The key contribution of this work is that it takes into account the nonuniform user distribution. Authors make use of a stochastic programming approach using mixed integer programming to model and solve the base station location problem from a BS power efficiency perspective. It is claimed a power reduction of at least 96% is possible with the proposed solution. However, the proposed techniques assume full knowledge of channel state information (CSI) at BS while neglecting effect of small scale fading and shadowing.

2) ENERGY SAVINGS THROUGH BS WITH PROPORTIONAL ENERGY MODEL

In [93] authors provide an analytical estimation of the energy savings that can be achieved for two BS models: a) On-Off BS energy model (current BS are more of this type) b) Proportional energy model where energy consumed is proportional to load in the cell. They use a QoS metric of delay, which is defined as inverse of throughput, which is further defined as an abstract function of distance from BS only, for analytical tractability. With this model, they present expressions for expected delay and variance of delay for given BS and user densities. These expressions are then used to analyze the BS densities and, hence, energy consumption for a given user density and QoS constraints. They also formally show the fact that lowest BS density for a given user density is possible with circular cells (as circle has largest area and thus largest number of users for a given distance allowed form the center of shape i.e. BS). The proof actually yields lower bound on the BS densities for any allowed topology of BSs (grid, hexagon, Poisson). The key inferences obtained are: 1) Poisson topology is less energy efficient than the regular topologies due to irregularly large distances from BS; 2) on/off model allows much more energy savings than the proportional model, advocating the use of system level techniques (compared to transmission power focused physical layer techniques that try to reduce variable energy consumptions on individual BSs). With proportional energy model (i.e. futuristic and unrealistic at the moment) the optimal energy saving model is not the one with lowest BS density. This advocates low power large number of small cells.

3) ENERGY SAVINGS BY SWITCHING ON/OFF BS

In [94], the authors present a methodology to calculate the energy savings by switching off BSs. They model energy consumption as a linear function of the number of BSs. Then, using the traffic profile for 24 hours, it is argued that as the traffic decreased by a factor X, a fraction X of the BSs can be shut down, and consequently, energy consumption will also be reduced by a factor X. Next, they remove the assumption that any BS can be shut down, and suggest that, in specific topologies, only certain BSs can be shut down to avoid coverage holes (e.g. in hexagon, six out of seven or three out of four BSs can be shut down). Similarly, authors identify the number of BSs for crossroad (urban street scenario with each cell having four neighbors) and Manhattan lay out. Note that the paper assumes omni directional antennas appointed in center of the cells and does not quantify the loss of coverage, capacity or takes into account the local user demands when shutting down BSs.

Authors in [93]–[95] present a scheme for energy management of base stations according to the network traffic that incorporates binary on/off activation or continuous cell zooming capabilities at the BSs. It is shown that noticeable energy savings can be achieved for low network traffic.

The authors in [96] present energy efficiency metrics and investigate the performance of different planning strategies of LTE networks in an empirical way. In [97], the authors propose to incorporate the on/off switching of BSs in the planning process itself. They first present a heuristic to have a minimum number of BSs. In this algorithm, first a Verona tessellation is established, then BSs are classified in feasible and infeasible set. A feasible set consists of BSs whose removal will not decrease the coverage below the threshold. This step is repeated until no BS can be removed without decreasing the coverage. To incorporate on/off switching, first the network is planned for the lowest traffic (this defines the infeasible set that cannot be switched off) and then it is planned for highest load. Turning on additional BS and finding their locations is done by repeating the same algorithm.

In [98], the authors use a detailed energy consumption model of the BS and definitions of site load factor to predict how much energy will be consumed to provide target capacity demands (100 Mbps in the paper). They take the energy consumption and capacity of existing HSPA+ in Finland (2008) as reference and compare it with that of LTE while considering the gains obtained in LTE by node level (energy/capacity) efficiency as well as by network level deployment strategies.

4) ENERGY SAVING THOUGH CELL SIZE ADAPTATION

In [99] authors present an analytical framework coupled with a simple mathematical traffic model, to investigate the potential energy savings that can be achieved by adaptively adjusting the cell size according to the spatial traffic variation.

The key idea is that, instead of having the same cell size throughout, areas with low traffic density can have larger cells compared to areas with high traffic density, resulting in energy and cost savings. The cell radius is calculated such that the cell includes maximum number of users while maintaining a threshold blocking probability resulting from underlying M/M/N/0 queuing model of traffic and geographically varying user density. The results show that energy savings increase as cell density decreases until a certain point, where the large transmission power overcomes the fixed power consumptions, and the energy savings start to diminish, reaching negative values. This work does not incorporate impact of interference, frequency reuse and sectorization on coverage.

In [56] the same authors present a modification of their work in [99] again for energy savings in LTE. Here instead of adapting the cell sizes in order to cope with spatial traffic variation, they propose dynamically adapting the number of sectors per site (three to two sectors per site) to cope with temporal variation of the sectors. They argue that though adaptive sectorization has been previously used for CDMA systems, it suffered from two short comings: 1) QoS constraints are not taken into consideration which makes the user terminals suffer from high blocking probability and low coverage once some sectors are off; 2) the transmission power of the BSs is assumed to be adjustable in a large dynamic range which is normally impractical in real systems because of the power amplifier and RF link. In order to overcome this shortcoming they suggest to increase the beam width and change the azimuth of two remaining sectors when the third is shut down in low traffic times. The QoS and power constraints are satisfied while saving 21% energy per BS using EARTH's measurement based power consumption model of the BS. Instead of full buffer traffic model, they use event triggered traffic model that is based on the continuous time Markov process. However, the results presented seem to be independent of the underlying traffic model.

5) ENERGY SAVING THROUGH NETWORK INFRASTRUCTURE SHARING

In [100], the authors exploit the observation that metropolitan areas are normally served by a few competing cellular network operators, providing 24/7 full coverage, each dimensioning its network according to peak traffic, but providing redundant resources when traffic is low. So provided that operators are willing to accept the competitor's subscribers as roaming customers, some energy can be saved. For the case of just two operators, the authors show that 20% or more energy can be saved, though the exact saving will depend on lot of parameters including operator policies, that remain to be investigated.

6) ENERGY SAVINGS THROUGH PROACTIVE APPROACH

The previously proposed approaches for switching BSs on and off are reactive, i.e., the network plan and BS sites are fixed and then reactive measures are implemented to decide which sites to switch on/off and which sites to reconfigure as network conditions change. To increase energy efficiency, an alternative approach would be to perform network planning as a function of the traffic load over large-scale time durations (hours and days) instead of planning only based on peak hour worst case traffic conditions. The process would then start by performing network planning based on the low traffic load conditions; this will give the set of BSs that need to be switched on at all times (during both high and low load conditions). The next step would be to perform network planning based on the higher traffic load conditions but with the sites obtained in the previous step fixed. The aim is to determine how many new sites to add and where to place them in order to support the increase in load; these are the sites to be additionally switched on in high load conditions. This approach, summarized in Fig. 3 is generic enough to account for any number of traffic load states depending on the large-scale traffic variations in the area of interest.

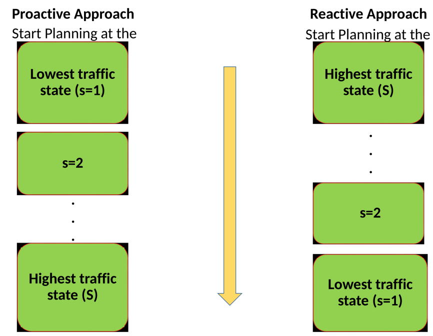

FIGURE 3. Green planning with BS on/off switching: proactive versus reactive approach.

A proactive approach proposing a modeling and optimisation framework for the planning of energy-aware wireless networks is introduced in [101]. The key idea is that energy awareness should be introduced at the planning stage in order to reach energy-efficient network operation. The authors formulate a joint planning and energy management problem, that aims to minimize a utility comprising a weighted sum of CAPEX and OPEX, including power consumption costs. By solving a binary linear program, the authors optimize BS positions, types, and configurations.

B. PLANNING WITH TRAFFIC UNCERTAINTY CONSIDERATIONS

Due to heavy traffic fluctuation over time cellular networks, operators often use peak hour traffic volume during network planning in order to avoid capacity bottlenecks [102]. A more efficient approach would be to consider the design of cellular networks under traffic uncertainty. For example, in [103] the authors suggest that better radio resource usage is possible by incorporating the time-varying traffic in the planning. They demonstrate that the user mobility can be similarly converted to the multi-period optimization problem. To this end, they formulate a simple one dimensional cell planning problem and demonstrate how it can be converted into a binary linear programming problem to look for the optimal solution.

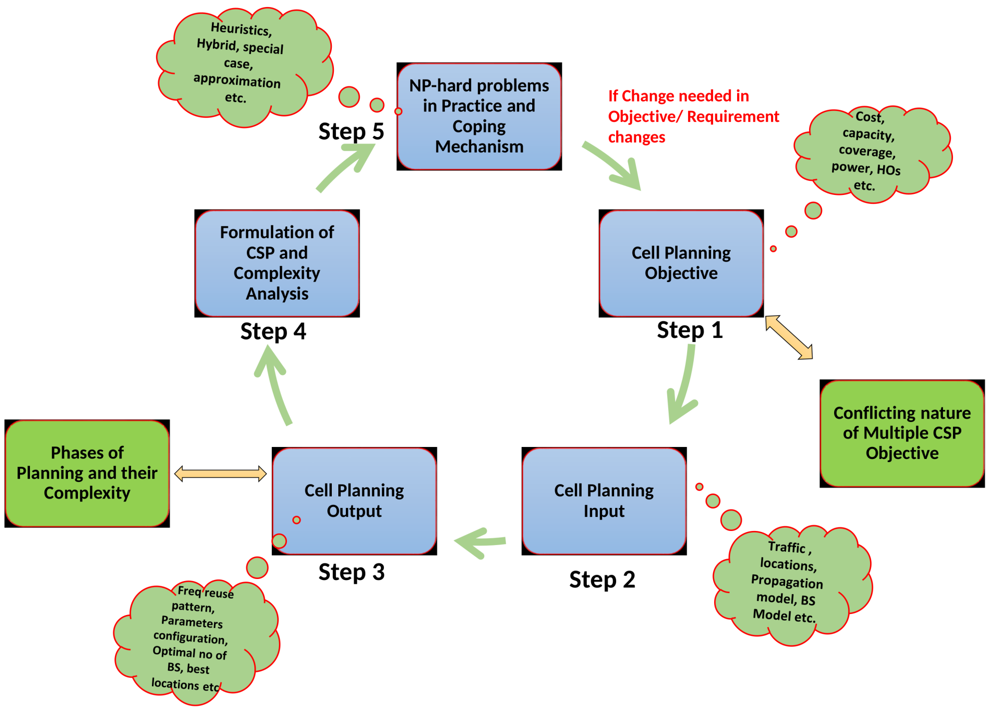

FIGURE 4. Typical phases of cellular network planning and parameters taken into consideration in each phase.

An overview of the most important techniques dealing with traffic uncertainty is presented in [104]. The literature that provides solutions for planning with traffic uncertainty can be divided in two categories discussed below.

1) MULTI-PERIOD NETWORK DESIGN

In multi-period (multi-hour) network design [105], an explicit set of demand matrices is given, and the network is designed in such a way that each of the demand matrices can be routed non-simultaneously within the installed capacities. In this context, authors in [106] introduce the concept of dominating demand matrices (i.e., D1 dominates D2 if every link capacity vector supporting D1 also supports D2). Instead of describing demand matrices explicitly, the authors in [107] consider the optimized routing of demands that may vary within a given prototype. For network design problems, this concept has mainly been applied using the hose model, a polyhedral demand uncertainty set which has been introduced in the context of virtual private networks (VPNs) [108]. More details and a good overview about methods to deal with network design under uncertainty are presented in [102], where robust optimization is advocated as a possible solution.

2) ROBUST OPTIMIZATION

Robust optimization was first considered by Soyster [109], and it aims at finding solutions that are feasible for all realizations of data in a given (bounded) uncertainty set. In robust optimisation a parameter Γ is used to control the price of robustness, the trade-off between the degree of uncertainty taken into account and the cost of this additional feature [110]. In [111] authors apply robust optimization to deal with demand uncertainty in cellular networks. Their robust optimization model offers for operators a trade- off between robustness and energy consumption by varying the robustness parameter. The complexity of problem is reduced by applying cutting planes. A case study is performed to compare the robust formulation to its deterministic counterpart and to conventional network planning. It is observed that energy savings are possible either by deploying less BSs or serving more users with the same number of BSs using the proposed robust optimization approach.

C. PLANNING WITH CoMP

In conventional CP, the BS coverage areas are controlled to minimize coverage overlap. However, when the BSs can coordinate to dynamically reduce interfere or balance loads, as in CoMP, standardised for LTE-Advanced (Release 11) and beyond, signal coverage overlap can be tolerated or even becomes desired. Thus, planning with CoMP becomes a very different problem compared to traditional planning discussed in previous sections, triggering some dedicated studies in recent years.

In [112] authors investigate the impact of coordinated multipoint (CoMP) transmission on cell planning parameters such as coverage, traffic, handover, and cost. To assess achievable coverage and capacities with and without CoMP, authors propose and use two ratios: Local to-uncooperative-plus-noise ratio (LUNR) and the local to-cooperative-ratio (LCR). Simulation results show that CoMP maximizes

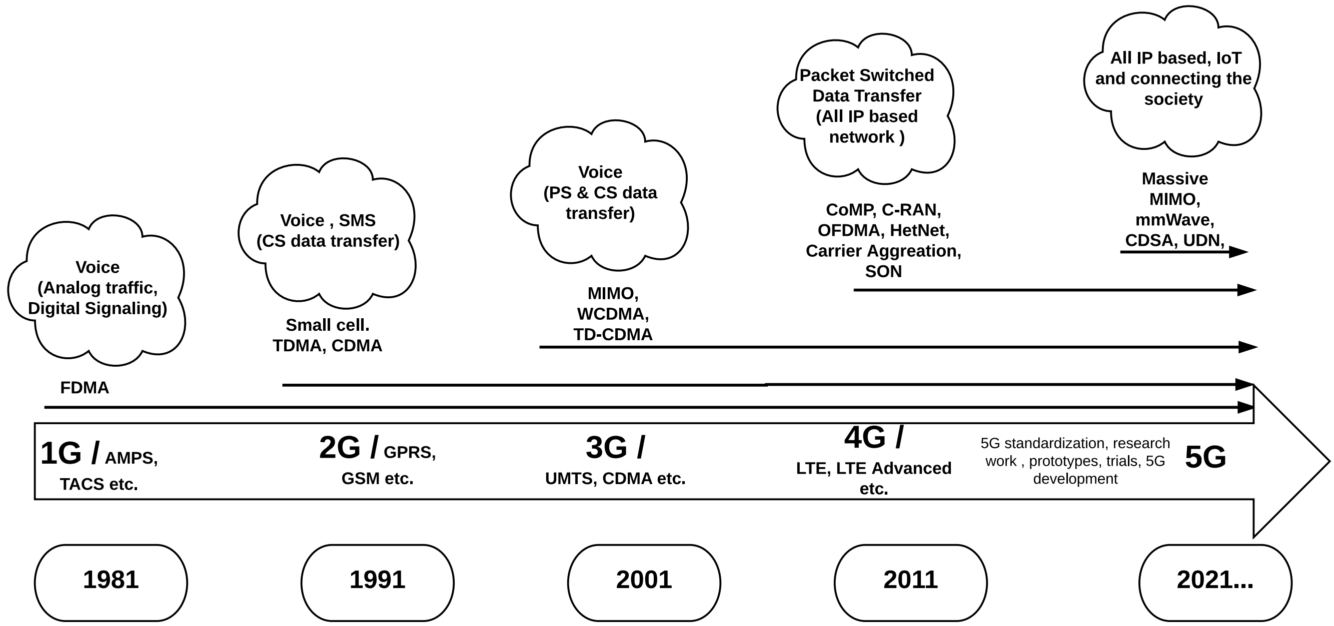

FIGURE 5. Brief history and time line of cellular technology generations. Different service offerings which became available in each generation are listed.

its gains over noncooperation (NC) in a network. However, NC may produce higher throughput in certain scenarios. Therefore, to avoid low system wide average throughput with CoMP, authors recommend a dynamic or semi-static switching between CoMP and NC called fractional base station cooperation. It is highlighted that because of interference from non-cooperating BSs, the gains of CoMP over NC are upper bounded and diminish at greater inter site distances due to noise. This encourages smaller cell sizes, higher transmit powers, and dynamic clustering of cooperative BSs. Cell planning with CoMP may require additional steps e.g. determination of cooperative and non-cooperative regions, LCR, and LUNR thresholds. Findings also show that gains of CoMP remain moderate, hence the complexity and cost incurred by CoMP should also remain moderate. The key limitation of this work is that each BS is assumed to be incapacitated. Therefore, user throughput solely determines the selection of the BS cluster in dynamic clustering, and load balancing remains an unexplored aspect.

In [113], the authors compare the energy saving potential of relay station (RS) and CoMP with single BS scenario while maintaining an average outage constraint. The impact of the traffic intensity and BS density are also investigated. Results show that traffic intensity can be divided into three classes: "coverage-limited" region, "energy-efficient" region, and "capacity-limited" region. The interesting finding is, as BS density goes higher, the energy-efficient region becomes larger, and the traffic load region where the cooperation schemes bring benefits becomes smaller. Furthermore, it is observed that RSs' energy cost needs to be designed as low as possible to get high performance, otherwise BS cooperation would be more favorable. The analytical models developed provide useful insights for green planning with CoMP

In [114], the authors extend their work in [113] for BS location and number planning taking into account CoMP, thus making the process of network planning more energy-aware. The optimal network planning problem is formulated as a mixed integer programming problem and an approximate solution is proposed using Lagrangian relaxation. Numerical results show that the overall energy consumption is decreased by over 20% compared with no cooperation while the system QoS is guaranteed. It is also observed that the low network density and traffic distribution asymmetry lead to higher energy efficiency gain.

V. NEW MODELING TOOLS FOR PLANNING FUTURE NETWORKS

In wake of new requirements and technologies discussed in previous section that are a becoming vital part of the CP paradigm, planning future networks call for corresponding evolution in modelling tools. In this section, we provide a brief overview of recent developments in models that have been proposed in literature to cope with new requirements in CP paradigm.

A. MODELING DYNAMIC CHANNEL VARIATIONS IN CP

In [115], the authors argue that conventional planning techniques rely on static propagation and interference models, which, although they take geographic information into account, they overlook the dynamic channel variations. To bridge the gap between static and dynamic planning the authors established a simple relationship that relates the static

SIR to its dynamic counterpart by a factor, which represents the influence of a fading environment. They use this new relation to study several aspects of a cellular system that reveal the physical implication of the static SIR in a dynamic operational environment. They also provide a simple method for evaluating the average outage probability. Finally, they determine the relationship between the conflicting requirements on the system capacity and on the minimum outage performance.

B. MODELING MULTI-CARRIER SYSTEMS IN CP

In [116], the authors argue that classic interference models used in CP have been limited to single carrier systems and present, an analytical method to assess effective SINR in multi-carrier systems operating over frequency selective channels. This extension of single carrier is achieved by expressing the link outage probability in terms of the statistics of the effective SINR. Two approximations for the link outage probability are obtained by considering Log-Normal and Gaussian assumptions for the derivation of the statistics of the exponential effective SINR. The SINR statistics are used to further assess the outage probability and thus obtain a simplified planning procedure for two cells interference scenario.

C. MODELING CELL LOAD IN CP

In [117], the authors introduced a method to analytically approximate cell load levels while planning OFDM networks by building on corresponding ideas in UMTS networks, namely load scaling and continuous traffic distributions. The idealized power control equations for UMTS are replaced by affine linear approximations of the adaptive modulation and coding (AMC) mechanism for OFDM which results in a different structure of the equations. They proposed a simple iteration for solving the equations numerically. The approach avoids time consuming snapshot simulations. The model is useful for automatic network planning and optimization as fast analytical capacity evaluation can build the foundation for various local search algorithms for improving network designs.

The authors in [118], extend their work in [119] and present a mathematical analysis of fundamental properties of the load coupling among cells for LTE system. They also develop and prove a sufficient and necessary condition for the solution's existence. Theoretical results for numerically approaching the solution or delivering a bounding interval are also presented. Finally, the application of the proposed system model for planning LTE network is presented. The analysis in [118] has been supported by theoretical proofs and numerical experiments and can serve as a basis for developing radio network planning and optimization strategies for LTE. Furthermore, the presented linearization and the bounding based optimization can potentially be used for more general convex optimization problems with similar properties. However, the analysis does not take into account interference dynamics specific to the LTE. In the SINR model of [118], interference is considered from all cells weighted with the load in those cells only. In other words if two interfering cells A and B are loaded 50% and 20%, the interference from these cells will be scaled by 0.5 and 0.2 when being received by the cell under consideration. This is more like CDMA where interference is independent of frequency reuse. In LTE presence of sub carrier allocation and scheduling means both cell A and B can still be interfering as long as weight is greater than zero depending on if that particular sub carrier is being reused or not. Furthermore, use of fractional frequency reuse adds another dimension to the interference modeling in LTE that is not considered in this paper.

D. MODELLING BS LOCATION RANDOMNESS IN CP

In [51], the authors argue that, although the problem of BS placement has been addressed with standard Voronoi partitioning, the standard Voronoi partition cannot be used in a scenario where the BSs have heterogeneous and anisotropic characteristics (directional antennas) or where geographical terrain is not planar two-dimensional. Therefore they suggest to use a generalization of the standard Voronoi partition replacing the usual distance measure with the concept of an abstract general function (named node function in the paper) associated with each site. The optimisation objective is formulated as a product of the node function with user density, integrated over the generalised Voronoi tessellation of each BS and then summed over all BSs. It is concluded that the solution will be to place each BS at the centroid of general tessellations (for which an abstract formula is given). But it is not clear from the paper whether the BS location or its corresponding tessellation will be decided first.

In [87] authors address BS location problem with a single objective of minimizing outage (evaluated by Monte Carlo simulations). Their case study assumes HSDPA system with pre-presence of fixed number of micro sites. For performance evaluation they make use of a planning tool and Monte Carlo method using real network data that includes: 3-D geolocations of the base stations, digital elevation map and digital clutter map data, antenna characteristics (pattern, tilt, and cable losses), total transmit power levels and spatial broadband traffic.

The proposed algorithm to address the problem, named SMART, consists of three simple steps.Greedy, Simulated Annealing (SA), Greedy with memory (Greedy Mem), and Simulated Annealing with memory (SAMem). It is shown that simulated Annealing and Greedy algorithms achieve almost the same optimal deployment when the number of optimization iterations is large and, the greedy algorithm converges much faster than simulated annealing algorithm when the number of optimization iterations is small. It is also shown that in practical network deployment scenario SMART outperforms previously proposed meta heuristics such as Deployment Formula Metric (DFM) based schemes proposed in [120] and [102].

E. MODELING INTERCELL INTERFERENCE

In [121], intercell interference modeling is performed in the uplink, while taking user scheduling into account. Scenarios with round robin, proportional fair, and maximum SINR are considered in the presence of various fading types. The obtained semi-analytical expressions are used to evaluate network performance metrics such as the outage probability, ergodic capacity, and average fairness numerically. The derived model can be useful as an input for radio network planning algorithms, in order to take inter-cell interference into account with user scheduling during the planning process.

Modeling intercell interference in the presence of uplink power control is investigated by the same authors in [122]. Fading is also incorporated in the models, along with basic scheduling assumptions. The expressions derived in [122] are then utilized to quantify numerically certain network performance metrics including average resource fairness, average reduction in power consumption, and ergodic capacity. Although the models of [121] and [122] are derived for the uplink, indications on their downlink extensions are described. Main limitation of these models stems from considering a single subcarrier. Generalization to the scenario of multiple OFDMA subcarriers with dynamic subcarrier allocation (as in the case of LTE scheduling) is a daunting task.

F. TOTAL COST OF OWNERSHIP MODELS

1) NET PRESENT VALUE (NPV)

In [123], authors use a planning tool to compare the performance of several algorithms, using the optimization objective of NPV. NPV takes into account expected revenues, CAPEX and OPEX (of BS as well as sectors) over a period of 6 years. Optimization is done in terms of BS locations, their numbers (out of a set of locations), and the number of sectors per BS. These results show the Tabu search performs well compared to other approaches though at the expense of additional execution time.

2) COST ANALYSIS OF BS AND RELAY STATION (RS)

In [124] authors do a cost analysis of joint BS and RS deployment in the context of LTE (they assume 2x2 MIMO). To this end, they use simple linear cost model that is sum of BS and RS densities weighted with their relative cost factors. Then via simulations they derive the curves for iso-capacities while varying BS and RS densities (per square kilometer). Hence, an increase in RS or BS density increases capacity in general, so by varying their densities reciprocally, the same capacities can be obtained. All such combinations of RS-BS densities that achieve same capacity (area spectral efficiency) make an iso-capacity curve. Then for a given iso-curve the point where the linear cost model defined earlier, is tangent to it, yields the least cost deployment for achieving that capacity. For this kind of static deployment, the optimal number of RSs is reported between 7 and 11. As after certain RS density, further increase in RS density can decrease the capacity (the same capacity cannot be maintained), therefore, iso-capacity curves also provide an upper bound on the RS density. Authors also analyze the impact of back haul distance on capacity-cost trade off of RS. They evaluate cost efficiency against number of RS per cell for various backhaul distances. Cost efficiency function proposed in [125] is used for this objective that is proportional to spectral efficiency (that is indirectly function of distance) and inversely proportional to the RS costs. The authors further investigate the impact of progressive deployment of RS. That is, unlike previous case where all RS are assumed to be deployed at once, authors assume RS might be deployed gradually e.g., one each year. To evaluate this scenario authors, use measure called ACSI (Average Customer Satisfaction Index, a term from economy), that quantifies user satisfaction after an upgrade relative to previous network construction. Results suggest that in terms of user satisfaction four relays are optimal.

In [126], the authors carry out cost/revenue analysis of WiMAX in presence of relays where revenue generated is modeled as function of capacity produced. They analyse the impact of location of RS, frequency reuse topologies and number of sectors on the cost/revenue optimization results show that trisector-ed BSs in topologies with relays enable the operators to achieve more profitable reuse configurations than with omnidirectional BSs and no relays. In [124], the authors investigate the possible energy gains of evolving a mobile network through a joint pico deployment and macro upgrade solution over a period of eight years. Besides the network energy consumption, energy efficiency in Mbps/kW is also analyzed. Outcomes of cost analysis in terms of total cost of ownership are shown for different deployment options considered. Using previous year of the evolution analysis, it is shown that deploying more pico sites reduces the energy consumption of the network, by a maximum of 30 percent. With regards to the energy efficiency, high deployment of pico sites allowed the network to carry 16 percent more traffic for the same amount of energy. This, however, results in an increase in operational costs.

VI. KEY CHALLENGES IN PLANING EMERGING HETEROGENEOUS AND ULTRA DENSE NETWORKS

In addition to the challenges and constraints already identified in section II and III in the context of classic and holistic cellular planning, modern day CP faces new challenges that stem from heterogeneity of the network, or more specifically advent of small cells. In this section, we discuss the challenges that are acting as Achilles heel for small cells and HetNets planning.

A. UL AND DL SYMMETRY

A peculiar feature of HetNets CP is pronounced uplink downlink asymmetry [127] that is generally neglected in most academic research studies on HetNets. Although this asymmetry exists in microcells only networks as well, the difference between uplink and downlink is potentially much larger in HetNets. The reason for this pronounced asymmetry in HetNets is as follows: In the downlink, transmit power disparities of 20dB+ exist between macrocells and small cells while this is not true for the uplink case wherein all transmitting UEs are roughly equal in terms of transmission power. In other words, in uplink direction, a small cell and a macrocell appear to be same for transmitting UE. Therefore, from UEs perspective, for the downlink, macrocells have very large coverage areas as compared to small cells while in uplink, coverage areas are roughly equal. As a result, cell association based on maximum received signal strength (Max-RSS) strategy may yield different optimal cells for uplink and downlink e.g., UEs that are connected to macrocell due to better downlink reception would likely be better off by associating with near-by small cells in the uplink. Not allowing this asymmetric association might lead to sub-optimal performance. On the other hand, allowing independent traffic sources in each direction raises new challenges for the core network and also for the UEs QoS e.g., cell edge UE may have poor SINR in one direction but not in the other. This different cell association in uplink/downlink will result in different interference models and resulting SINRs in the two links e.g., UEs sharing same BS will be orthogonal to one another on the downlink while they may interfere with each other on the uplink if they are transmitting to different base stations. This calls for new two-way channel models to be investigated and incorporated in the CP as the channel gains and SINR in the two directions may be almost uncorrelated especially if they are routed from different base stations.

B. BACKHAUL

Ultra-dense deployment of small cells will create additional challenges for the transport of backhaul traffic. Current [128] research studies on HetNets utilize small number of small cells to improve SINR of wireless links in limited hotspot areas wherein a relatively small backhaul traffic originating from small cells can be forwarded into the core network through conventional backhaul links of cellular networks. These studies [128] focus on gains of the wireless front haul and neglect any possible backhaul bottlenecks. This assumption is generally correct for well-planned conventional macrocell only cellular networks. But this assumption breaks down for HetNets where small cells are ultra-densely deployed, as in such deployments it may become a key problem to forward massive traffic into the core network through existing backhauls. It is now believed that the full benefits of dense HetNets can be realized only if they are supported by the careful backhaul planning [129].

Femtocells deployed in homes by the subscribers generally utilize digital subscriber line (DSL) broadband connection for the backhaul that can quickly become bottleneck particularly in the Uplink [127]. IP traffic through traditional internet service providers (ISPs) is used to connect the femtocells with the core network. This demands high QoS requirements from broadband connections. Several things need to be considered like whether the broadband support QoS or traffic prioritization or is the connection throttled or traffic-shaped by ISP. The probable high latency in broadband backhaul can pose serious problems in coordination of resource allocation or handoffs information with other cells. Utilizing untrusted IP network for backhaul poses serious security issues as well. Macrocells enjoy commercial grade security which is absent in small cells as they are in direct reach of the subscribers. As a result, small cells and public IP networks can be utilized to launch distributed attack on cellular network. Security planning challenges become another challenge for these unplanned networks.

C. CP FOR HetNets WITH ENERGY EFFICIENCY CONSTRAINTS

Energy Efficiency (EE) is emerging as one of the main challenges in rolling out HetNets. In this section we describe what makes EE a key constraint to be considered while planning HetNet deployments.

1) PLANNING FOR EE FROM NETWORK PERSPECTIVE

First factor that makes EE more significant in HetNet planning compared to old macro cell only network planning is the sheer increase in number of cells. A large portion of the energy dissipated in a cellular system is actually consumed at the base stations (BSs) [130]. Although the small cells have a relatively lower power consumption profile, ultra-dense deployment can lead to high aggregated energy consumption. From EE perspective, small cells are beneficial only when they are deployed in ideal locations where data requirements are high or macro cell performance is low [131]. On the contrary, small cells in control of subscribers like femtocells may not be beneficial in terms of aggregate EE as these small cells are operational at all times of the day. Even in the absence of users in their coverage, a substantial amount of circuit energy is drawn by these nodes. Therefore, switching certain base stations off in light traffic conditions, is an efficient technique to save energy in wireless networks [130]. However, to implement such technique CP has to be carried out in way that switching off those certain cells does not create coverage holes. Another alternative is to put certain base stations into the sleep (dormant) mode. However, such dormant cells need to be preemptively activated when user devices are moving into their coverage and new capacity is needed. For this to happen, the network must be able to wake the dormant small cell before handing over traffic to the dormant cell. This requires CP such that all sleep enabled cells have some sort of fast signaling connectivity with neighboring cells either on front haul or back haul.

One simpler approach is shutting down almost all the modules of a small cell based on a fixed timer configured based on statistical traffic cycle [131]. However, implementation of such scheme requires carrying out what is called multimodal CP which in addition to spatial variation in traffic has to consider temporal variation in traffic. This kind of CP is even more challenging and has yet to be investigated. Another newly conceived constraint related to EE is CP where cell sites rely on renewable energy source. In this type of planning in addition to spatio temporal variation of traffic, spatio temporal availability of the renewable energy source has to be taken into account in the CP optimization problem [132].

2) PLANNING FOR EE FROM UE PERSPECTIVE

From UE's perspective, energy consumption becomes an issue in ultra-dense deployment especially if small cells utilize separate frequency bands/RAT. UE's must periodically scan for nearby small cells for traffic offloading opportunity that can result in significant energy consumption for the UE. Therefore, energy efficient discovery of small cells becomes a problem in carrier frequency separated deployment and a balanced inter-frequency small cell discovery (ISCD) interval needs to be optimized. On one hand, low ISCD periodicity (i.e. high scanning frequency) can result in increased small cell offloading opportunity, thus enhancing the capacity and coverage. However, this can also lead to higher UE power consumption due to the high scanning frequency and lower transmit power of small cells. On the other hand, high ISCD periodicity (i.e. low scanning frequency) can lead to the UE missing small cell off-loading opportunity, thus resulting in a potential decrease in capacity.

In a recent study on ISCD [133], it has been shown that for given cell density and UE mobility there exists an optimal ISCD frequency in terms of EE. As UE battery life is already a bottle neck in era of smart phones, this finding needs to be incorporated in CP for HetNets to determine optimal cell densities, not only from capacity perspective but EE perspective.

CP for emerging HetNets, need to take all of the aforementioned EE constraints into account, which make it a drastically different problem compared to the ones studied extensively in literature for macro cell planning with focus on coverage and capacity only.

D. CP WITH DEPLOYMENT LOCATION CONSTRAINTS IN HetNets

1) BS PLACEMENT OPTIMIZATION IN HetNets

Planning of small cells was investigated in [134] assuming macrocell locations are fixed. In [59], the BS placement optimization was performed jointly for macrocell and small cell BSs in a non uniform user density scenario. A heuristic approach based on simulated annealing was adopted, taking into account intercell interference, dynamic resource management, and joint uplink/downlink performance. The authors demonstrate the efficiency of HetNet planning in a hotspot scenario, shown in Fig. 7, simulated using a Gaussian user distribution assuming four macrocell BSs and 64 small cell BSs. In fact, the simulated annealing approach move the small cells from their rectangular grid positions to a Gaussian deployment, following that of the user distribution. Such hotspot areas are characterised with temporary traffic surges, such as football stadiums, where the user density is very high around a football match and very low otherwise. For this reason, the HetNet deployment scenario is convenient, since small cells can be switched on and off according to the varying traffic.

2) SENSITIVITY TO SMALL CELL DEPLOYMENT LOCATION IN HetNets

While planning, conventional cellular networks consisting only the macro cells, even if the macro cells are not deployed at the ideal location but somewhat near the optimal location, the larger radius of the macro cell, and the ability to tweak antenna tilts and azimuths compensates the difference between optimal and the actual location. This tolerance for difference between actual and optimal location in case of small cells decreases due to their small coverage areas and relatively inflexible antennas. As a result, contrary to popular belief, physical location of the small cell requires more precise engineering than macro cells. A slight difference between the optimal and actual location of small cell due to the physical limitations or real estate constraints can actually degrade network performance as small cell location can have large impact on interference pattern and mobility related performance. Therefore, it becomes imperative to make optimal decisions in HetNet planning for small cell locations to achieve efficient performance of the network.

3) CP WITH RS POSITIONING CONSIDERATIONS IN HetNets

Most studies that investigate RS based deployments of HetNets generally assume RS locations to be prefixed by some arbitrary criterion. On the contrary to harness the full advantage of RS its location is expected to be more impromptu than BS and hence the impact of location of RS on the performance of the system has to be investigated. In [135], the authors address this challenge by building on previous work in [136] and present an analytical study of RS positioning. The impact of RS location selection and cell selection on the system performance are evaluated in a single interferer scenario. Closed-form expressions for the link SIR, link rate, and end-to-end user rate distributions are obtained. Similarly, in [137] authors have highlighted that existing theoretical analysis on RS performance assessment has primarily focused on Gaussian relay channels, and the analysis of interference-limited relay deployment has been confined to simulation based approaches. In their paper, they take the initiative to provide analytical closed form expression to asses system capacity as function of the optimal location and number of relays, and resource sharing between relay and base-stations. The paper shows that the optimal deployment parameters are pre-dominantly a function of the saturation capacity, path loss exponent and transmit powers.

E. CP WITH OFF-LOADING CONSIDERATIONS IN HetNets

In [134], the authors attempt to optimally plan small cell locations for offloading traffic from macro to small cells in a HetNet LTE cellular network. The challenge is complicated further by interference coupling between cell loads in a non-linear manner. A search algorithm leading to near-optimal

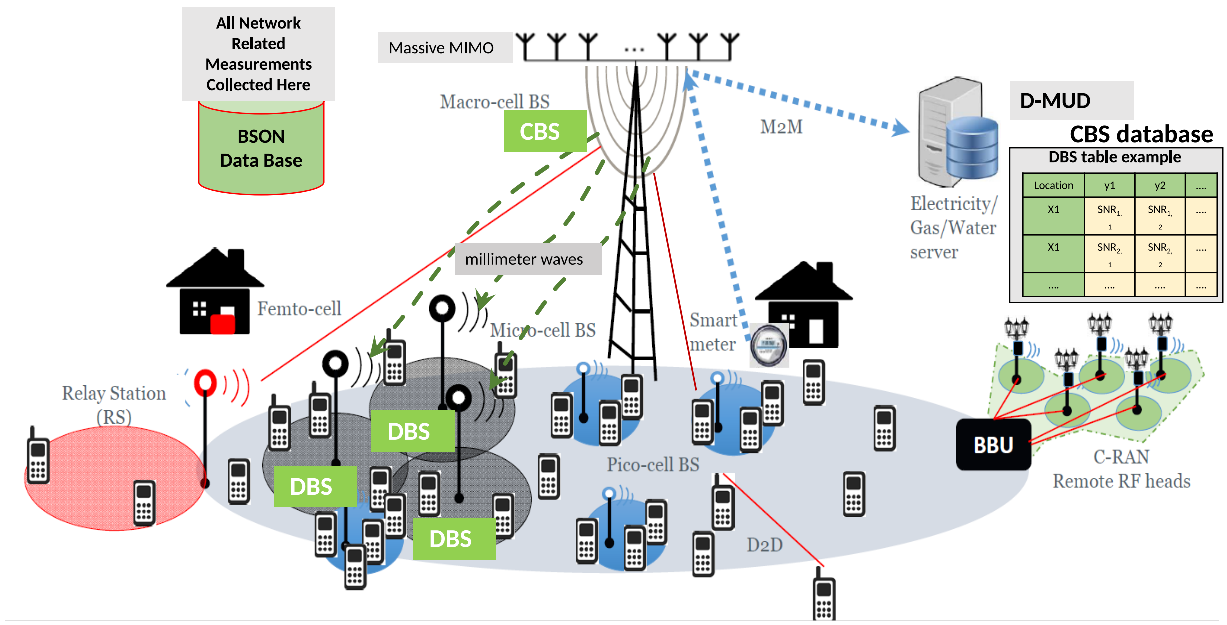

FIGURE 6. 5G networks are envisioned to be multi-layered and multi-RAT consisting of macrocells, micro-cells, pico-cells, and relays, with device to device communication, mmWave, massive MIMO and cloud-RAN.

solutions is proposed. Its objective is to select, from a set of candidate locations, up to a given number of small cell sites in a deployed macro cell network. In addition, an approach for numerically constructing a tight linear approximation is proposed, in order to enable the use of mixed integer linear programming to gauge optimality.

F. PLANNING WITH MULTI-RAT

Beyond these early pioneering works on HetNets planning discussed in previous subsections, detailed network planning in the presence of HetNets still needs a more thorough and detailed investigation. Other more complex scenarios, such as planning a network in the presence of femtocells [138], that can be considered as small cells deployed without operator control, make the problem more challenging. Another relevant scenario is the planning with multiple radio access technologies (RAT), where other technologies, such as WiFi, can be used to offload cellular traffic. This multi-RAT operation can be uncontrolled, similar to the case of femtocell deployments, or operator-controlled, where the mobile operators deploy WiFi access points to offload part of the cellular traffic.

VII. CHALLENGES AND OPPORTUNITIES IN PLANNING FUTURE CELLULAR NETWORKS: WHAT IT WILL TAKE TO PLAN 5G NETWORK?

In the wake of 5G, CP is faced with numerous challenges, some evolving from 4G such as HetNets, carrier aggregation, inter-cell interference coordination and CoMP, others that are characteristic of 5G, such as the C-RAN, D2D, M2M communication, mmWave and Massive MIMO based deployments. Fig. 6 shows a typical 5G deployment consisting of a macro-cell, under-laid with a heterogeneous mix of small cells including: micro-cells, pico-cells, RS, remote radio heads (RRH), femtocells, CDSA, D-MUD and D2D. In this section, we discuss how the adaptation of these new technologies in 5G and beyond may affect the CP paradigm.

A. PLANNING WITH CLOUD-RAN (C-RAN)

Dense deployment of small cells requires centralised coordination to avoid inter-cell interference and provide intelligent resource allocation in response to spatio temporally varying traffic. The Cloud or Centralised RAN architecture, is thus considered a prime enabler to ultra-dense networks as it allows the required coordination. Basically, the C-RAN consists of breaking of the traditional eNB functions and migrating them towards the centralised processor. As shown in Fig. 6, the centralised node holds the base band functions, and is called the base band unit (BBU). The BBU is then connected to a many low complexity access point, which often only consist of radio and analog/digital functions, called the remote radio head (RRH). Hence, from a radio point of view, the C-RAN has solved the radio access bottleneck in a flexible, scalable and adaptive manner due to the ease of resource allocation among RRHs and addition or relocation of RRHs. Moreover, owing to their low complexity, RRHs are low cost, robust, and small in size, thus, result in the reduction of the CAPEX of deployment and the OPEX of maintenance and premises rent. Nonetheless, a new bottleneck is born with the C-RAN; that is the link connecting the RRH to the BBU, referred to as fronthaul.

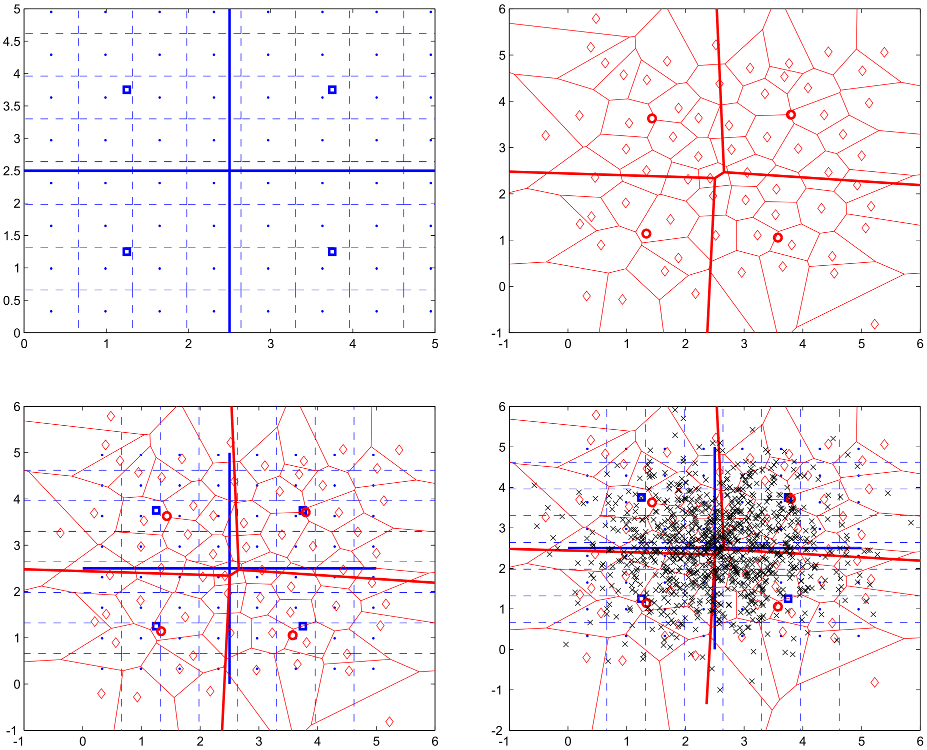

FIGURE 7. BS deployment (HetNet: 4 macro BSs and 64 small cell BSs) for a Gaussian user distribution over 25 km^2 with a density of 40 users/km^2. Upper left: Initial deployment. Upper right: Optimized deployment (Simulated Annealing). Lower left: Superposition of the two deployments. Lower right: deployments with the distribution of user terminals in the network (black 'x"s).

The emerging C-RAN presents a categoric shift in both coverage and capacity planning, while extending the CP targets to the front haul. Coverage planning becomes cell-less, since adaptive and variable sets of RRHs would now form virtual cells, replacing the traditional base stations. Capacity planning also evolves from being cell-centric to BBU-centric or User centric, consequently, improving resource usage efficiency. As for the front haul, it has henceforth become an integral part of the virtual C-RAN cell, hence, the corresponding CP approach.

1) CP WITH RADIO GAIN AND FRONT-HAUL COST CONSTRAINTS IN C-RAN

Authors in [139] look at the trade-off between the radio gains and front haul cost for different levels of function migration, otherwise referred to as functional split. Essentially, the C-RAN CP requires joint planning of the radio sites and the front-haul, as in [140], which looks at finding the RRH locations with a passive optical fiber network (PON) for the fronthaul. The topic of the paper is, thus, the infrastructure deployment and layout planning problem under the C-RAN architecture. It is formulated as a generic integer linear programming (ILP) model which aims at minimising the deployment cost, by identifying the locations of RRHs and optical wavelength division multiplexers (WDM) and their corresponding association relations, with the constraint of satisfying the coverage requirement. The optimisation framework proves to be solvable and scalable as validated through various case studies. Moreover, the results show significant gains, when CoMP is used in the C-RAN architecture, in terms of higher capacity and reliability at lower cost.

2) CP WITH JOINT RESOURCE MANAGEMENT CONSTRAINTS IN C-RAN

Authors in [141] look a novel framework for joint resource management in a HetNet with multi-RAT and C-RAN. The framework consists of categorizing various functionalities of the radio access and the fronthaul (PON-based) depending on the time requirement to conduct the management actions. Self-organization and cognitive capabilities are also incorporated in the framework, which could be applied to various phases of the network's life such as planning, deployment, optimization, etc.

Authors in [142] also analyze the system capacity in a C-RAN architecture, comparing two different CoMP options with factional frequency reuse (FFR). A multiple input single output (single user) scenario is generated using joint transmission, and MIMO scenario (two users) is created with beamforming, both assuming two RRHs. The authors demonstrate an extra 6dB downlink capacity gain with coordinated beamforming, however, at the expense of additional computational power for user pairing and selection.

3) CP WITH USER-CENTRIC CELLS IN C-RAN

Authors in [143] revamp the common understanding of cellular structure by proposing a user-centric virtual cell, formed by the user at its center and a cluster of remote radio head (RRH) around it. Tight cooperation between these cells is possible, allowing efficient power allocation. The system interference is first modeled based on the mean and variance, then the results are applied to find the optimum cell radius that would maximize the downlink system capacity.

B. PLANNING IN THE PRESENCE OF M2M AND IoT

The introduction of smart cities and the IoT in which homes, smart vehicles, sensing systems, and mundane objects are endowed with high-speed machine-to-machine (M2M) communication capabilities is seen as the major technological challenge for the next decade [144]. While traditional M2M communications has relied on short-range technologies such as Bluetooth or ZigBee, moving toward large-scale M2M smart cities requires broader interconnection and communications among machine type devices which is made possible by enabling M2M communications over the reliable cellular network infrastructure [145]. However, realizing this vision is contingent upon transforming the cellular infrastructure into a scalable and efficient system capable of sustaining the diverse challenges of M2M communications [144]–[146].

Consequently, novel cellular planning approaches are required for networks with M2M services. According to the authors' knowledge, the most of recent work on CP does not account for M2M/IoT deployments. M2M/IoT service service characteristics are typically different from traditional human-to-human (H2H) services; M2M/IoT services are distinguished in most applications by low cost, low mobility, delay tolerance (except urgent security and health cases), large number of devices (e.g., up to 30,000 smart meters per cell), generally small and infrequent data transmission, as defined in [147]. These variations raise pertinent questions about cellular network planning with M2M/IoT services, and pose diverse challenges in accommodating both M2M/IoT and human to human (H2H) traffic classes fairly and efficiently. A recent study item by 3GPP [148] gives an insight on M2M/IoT device characteristics that are relevant to network planning, such as: single receive antenna, reduced transmit power, reduced peak data rate of up to 1 Mbps, device noise figure of 9 dB, etc. In addition to the device characteristic constraints, M2M/IoT mobility behaviors and quality of service requirements vary greatly per M2M/IoT application. M2M/IoT over cellular networks allows to cost effectively and efficiently connect heterogeneous M2M/IoT devices such as vehicles, smart meters, sensors, and surveillance apparatus, among others. However, reaping the benefits of M2M/IoT deployment over cellular requires overcoming major technical challenges in CP while taking into account architectural, operational, and economic perspectives [144], [149]–[151].

1) ARCHITECTURE LEVEL CHALLENGES

At the architectural level, M2M communications over cellular systems will significantly increase the heterogeneity of the wireless landscape in terms of device types and traffic classes. Modeling, analyzing, and managing such heterogeneous M2M systems and incorporating those models in CP thus becomes essential [152]. On the one hand, planning of new nodes, such as M2M traffic aggregators or road side units must be deployed and integrated into the CP which traditionally has been focused on BS planning. On the other hand, expanding the network via new base stations or optimized frequency planning must now account for scattered, large-scale, and diverse M2M traffic, that must seamlessly co-exist with conventional cellular services.

2) OPERATIONAL CHALLENGES

From an operational point of view, the introduction of M2M will dramatically increase the amount of data circulating in the cellular network. This data explosion will naturally strain the already resource-constrained infrastructure, thus introducing major challenges for resource management and optimization [144], [147], [149]–[151] that include: 1) managing constrained resources (power, time, frequency) for large data volumes, 2) handling excessive M2M-service dependent signaling that modern-day cellular systems are not designed to sustain, 3) meeting heterogeneous QoS constraints of diverse M2M services with little disruption to legacy human to human (H2H) communications, 4) handling the high-speed mobility and dynamics incurred by vehicular M2M traffic, and 5) maintaining a self-organizing and cost-effective operation in a dense and heterogeneous network environment. In addition to these constraints, planning of an M2M enabling cellular network will also require consideration of features such as data aggregation for sensor-based M2M traffic, direct vehicle-to-vehicle communications, and coordinated communication for distributed M2M services.

3) ECONOMICS AND BUSINESS CHALLENGES

At the economic level, CP for M2M services requires new business models that allow to seize the various opportunities brought forward by M2M communications [152], [153]. One key issue is to develop market models for inclusion in CP to analyze the interactions between several key players in the M2M domain that include mobile operators, M2M providers, and possibly governmental agencies [153]. In addition, pricing will constitute an important factor in CP for M2M. There is a need to develop and incorporate models for pricing mechanisms for M2M services. The key challenge here is that, pricing is largely intertwined with both network planning and resource management, as it is a key determinant of how and which cellular resources are being used by the different M2M providers.

C. CP IN THE PRESENCE OF D2D

D2D communications have been receiving significant research attention recently, due to their planned incorporation in Release 12 of LTE-Advanced (LTE-A). D2D communications in LTE-A would allow a device to use the cellular spectrum in order to be connected directly to another device. Consequently, a transfer of large data amounts (e.g. multimedia) can occur through a direct connection over short distances. This short range (SR) D2D transfer permits to offload some traffic from the cellular network, since it does not need to use the network itself. D2D communications can take place in one of three following modes:

1) D2D terminals can use dedicated resources assigned to them by the cellular network.
2) They can reuse the same resources of the cellular network.
3) They can form an underlay network [154]–[156].

In the following we discuss the CP challenges, introduction of D2D puts forth.

1) FAIRNESS AND INTERFERENCE MINIMIZATION

In an underlay scenario the main challenge with D2D communication is to keep the interference with the primary cellular network at tolerable levels. In [154], a one-to-one reuse problem is adopted, where each D2D connection has to reuse a channel used by a cellular connection. The solution minimizing the interference is obtained using the Hungarian Algorithm. A joint scheduling and resource allocation scheme is proposed in [155] for a similar underlay scenario. The authors of [155] investigate a tradeoff between system throughput and user fairness through the definition of a fairness coefficient. In [156], the interference between the D2D network and the cellular network is controlled by keeping a minimum distance between a cellular transmitter and D2D receiver. Round Robin scheduling is used to ensure fairness. However, in the approach of [156], an extra overhead is incurred since devices need to report their location information to the BS.

2) SPATIAL REUSE

Maximizing the spatial reuse while us- ing D2D communications on dedicated or same channel is something that has to be considered in CP. From network perspective, small cells offer more aggressive spatial reuse. But with D2D in picture problem becomes twisted by the fact that small cells are more vulnerable to D2D interference due to close proximity to device and low power, compared to macro cells. Therefore, D2D considerate CP for HetNets is challenging problem which remains to be investigated. For pure macro cell based networks reuse maximization problem is relatively more tractable. For example, a D2D resource allocation scheme for maximizing spatial reuse is proposed in [157], where the BS allocates D2D channels in a relatively slow time scale whereas the mobile terminals (MTs) involved in a direct D2D link can adjust the modulation and coding scheme (MCS) level in a relatively fast time scale.

3) CP WITH D2D AND M2M INTERPLAY CONSIDERATIONS

Network planning with D2D considerations alone is an open problem. This problem is further aggravated by the fact that M2M and D2D can have opposing effects on the network operation: M2M adds additional traffic to be supported by the cellular network, whereas D2D communications could offload some traffic from the cellular BSs. Thus, the network planning process needs also to take into account the interplay between D2D and M2M when they coexist in future cellular networks.

D. CP WITH CONTROL DATA SEPARATION ARCHITECTURE (CDSA)

An undesired byproduct of inevitable densification of future network is that there will be a huge signaling overload specifically during scenarios such as mobility and handover, if conventional signaling procedures are used. As investigated recently [158], a better way to plan for next generation 5G networks is to leverage CDSA. The main idea of the CDSA originates from the fact that only a small amount of signaling is required to enable ubiquitous coverage [159]. On the other hand, data transmission and its related signaling are needed on demand only when there are active user equipment (UE). This calls for a two-layer RAN architecture with a logical separation between:

- Network access and data transmission functionalities.
- Idle mode and active mode.
- Cell-specific/broadcast-type and UE-specific/unicast-type signaling.

In CDSA, a continuous and reliable coverage layer will be provided by control base station (CBS) at low frequency bands, where the large footprint ensures robust connectivity and mobility. The data plane (DP) is supported by flexible, adaptive, high capacity and energy efficient data base stations (DBSs) that provide data transmission along with the necessary signaling. As shown conceptually in Fig. 6, all UEs are anchored to the CBS, while active UEs are associated with both the CBS and the DBS in a dual connection mode [160]. CDSA offers a range of benefits such as better energy efficiency and system capacity and resource efficient support for mobility as well as M2M/IoT. However, the concept of CDSA is still in early stages and specifying functionalities of each plane is not trivial and is an open research problem. Examples of recent work that investigate this problem include [161], [158].

Expectedly, CDSA will have a major impact on the way cellular systems are planned. Specifically, it will expand the dimensions of the solution space of the CP by requiring planning of two different nodes DBS and CBS. Additionally, the distinct requirements of coverage and capacity of both DBS and CBS, delay between DBS and CBS, proximity between DBS and CBS, will impose new constraints on DBS and CBS location, power, coverage and capacity planning.

E. CP WITH mmWave BASED CELLS

Promise of mmWave stems from two factors: 1) abundance of spectrum, 2) noise limited operation regime thanks to high propagation loss and thus short range which opens the doors for even denser deployment. However, the very same blessing associated with mmWave, i.e., a short range, is also a curse. Unlike sub 5GHz deployments where densification is a choice, and level of densification can be tailored to meet capacity demands as coverage is not a major constraint due to low propagation losses, mmWave based deployments have to be extremely dense to provide a barely acceptable level of continuous coverage. Several recent studies show [43] a maximum range of 100-200m, in line of sight (LoS) conditions. This means if mmWaves based small cells are adapted in future networks they will have to be adapted as complimentary source of capacity, while primary source of coverage has to be a high frequency (HF) based deployment of macro and small cells. One possibility to do so, as recently proposed in [158] is to exploit CDSA such that mmWave is used for DBS and HF spectrum is used for CBS. This immediately makes the pandora of CP challenges discussed in context of HetNets as well as CDSA relevant to mmWave based CP. Additional new challenge will be incorporation of the fact that mmWave requires LoS, and mmWave based cells might offer highly directive antennas. This may require planning of DBS with the consideration of cell less deployment where coverage will follow users in the form of narrowly focused beams, as such no fixed cell foot prints will exist. This will require a major shift from traditional CP, where cell foot print and their ability to cover the whole area of interest plays the pivotal role in CP problem formulations and solution search.

F. CP WITH MASSIVE MIMO BASED CELLS

The major challenge in planning massive MIMO enabled base stations stems from the fact that the large antenna array gain complicates the max-RSS based cell association problem further in HetNets. As per [35], even when macro BS reduces its power to same level as that of small cell, the user has a higher probability to still get connected to the macro due to a large gain of massive MIMO macrocell. As a result, this gain can force the massive MIMO macrocell to carry most of the data traffic in HetNets, resulting in a significant load imbalance between the macrocells and small cell. The second main challenge in CP with massive MIMO is accommodation of two conflicting objectives:

Two Conflicting Objective in CP With Massive MIMO: There are two opposing forces at work, while deploying massive MIMO in HetNets. Massive MIMO lends its gain from channel diversity. However, in ultra-dense small cell deployment, small cells owing to their small coverage area may result in large spatial correlation of the channels limiting degrees of freedom available to Massive MIMO thereby un- dermining their gain. As networks becomes denser the number of active users per cell will decrease and the need for massive MIMO may decrease. Other factors such as cost, energy and backhaul need to be taken into account for association/offloading decisions.

Therefore, planning of future networks have to strike a balance that might exist in the two extremes of all macro cells with massive MIMO, and HetNets with massive MIMO only on cells with size above a certain threshold. Investigation of this optimal cell size while taking into account the user density, number of antennas per site, channel types and TCO is a research problem that demands extensive study.

VIII. CONCLUSIONS

A plethora of new technologies need to be incorporated in future cellular networks to meet unprecedented traffic demands and to provide newly conceived services such as IoT/M2M. These technologies include HetNets, CoMP, D2D, CDSA, mmWave and massive MIMO. The adaptation of these technologies means that one of the oldest but still not fully matured research area in cellular networks, i.e. cell planning needs a major paradigm shift. This article serves as comprehensive reference to kick start the research in CP for 5G networks for both academic and industry based researchers.

To this end, in this article we provided timely analysis of this paradigm shift needed in CP. We start with a tutorial of CP to identify the input and outputs of typical CP problem and to characterize its computational complexity. We then provide concise recap of past attempts on different variants of classic CP problem. This is followed by a analysis of recent developments in CP to incorporate e.g., EE, traffic uncertainties and CoMP. We then provide an overview of recent advancements made in modelling tools to make CP problem more tractable and/or realistic. We then provide a comprehensive analysis of challenges that cellular industry still faces in planning HetNets, along with promising approaches to address these challenges. We conclude this article with a detailed discussion challenges and research opportunities in CP for 5G and beyond that stem from introduction of IoT/M2M, D2D, CDSA, mmWave and massive MIMO.

ACKNOWLEDGMENT

Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect those of the National Science Foundation.

REFERENCES

[1] Z. Fluhr and E. Nussbaum, "Switching plan for a cellular mobile telephone system," *IEEE Trans. Commun.*, vol. 21, no. 11, pp. 1281–1286, Nov. 1973.

[2] J. D. Wells, "Cellular system design using the expansion cell layout method," *IEEE Trans. Veh. Technol.*, vol. 33, no. 2, pp. 58–66, May 1984.

[3] A. Gamst, E. G. Zinn, R. Beck, and R. Simon, "Cellular radio network planning," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 1, no. 2, pp. 8–11, Feb. 1986.

[4] S. T. S. Chia, "Design and optimisation for cellular access network," *Electron. Commun. Eng. J.*, vol. 8, no. 6, pp. 269–277, Dec. 1996.

[5] M. Frullone, G. Riva, P. Grazioso, and G. Falciasecca, "Advanced planning criteria for cellular systems," *IEEE Pers. Commun.*, vol. 3, no. 6, pp. 10–15, Dec. 1996.

[6] E. Amaldi, A. Capone, and F. Malucelli, "Planning UMTS base station location: Optimization models with power control and algorithms," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 939–952, Sep. 2003.

[7] F. Perez-Fontan and J. M. Hernando Rabanos, "Educational cellular radio network planning software tool," *IEEE Trans. Edu.*, vol. 41, no. 3, pp. 203–215, Aug. 1998.

[8] K. Lieska, E. Laitinen, and J. Lahteenmaki, "Radio coverage optimization with genetic algorithms," in *Proc. 9th IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, vol. 1. Sep. 1998, pp. 318–322.

[9] S. Dehghan, D. Lister, R. Owen, and P. Jones, "W-CDMA capacity and planning issues," *Electron. Commun. Eng. J.*, vol. 12, no. 3, pp. 101–118, Jun. 2000.

[10] A. Ahmad, "A CDMA network architecture using optimized sectoring," *IEEE Trans. Veh. Technol.*, vol. 51, no. 3, pp. 404–410, May 2002.

[11] A. Imran, A. Zoha, and A. Abu-Dayya, "Challenges in 5G: How to empower SON with big data for enabling 5G," *IEEE Netw.*, vol. 28, no. 6, pp. 27–33, Nov. 2014.

[12] R. G. Akl, M. V. Hegde, M. Naraghi-Pour, and P. S. Min, "Multicell CDMA network design," *IEEE Trans. Veh. Technol.*, vol. 50, no. 3, pp. 711–722, May 2001.

[13] P. R. Gould, "Radio planning of third generation networks in urban areas," in *Proc. 3rd Int. Conf. 3G Mobile Commun. Technol.*, May 2002, pp. 64–68.

[14] S. Hanly and R. Mathar, "On the optimal base-station density for CDMA cellular networks," *IEEE Trans. Commun.*, vol. 50, no. 8, pp. 1274–1281, Aug. 2002.

[15] S. Hurley, "Planning effective cellular mobile radio networks," *IEEE Trans. Veh. Technol.*, vol. 51, no. 2, pp. 243–253, Mar. 2002.

[16] J. Munyaneza, A. Kurien, and B. V. Wyk, "Optimization of antenna placement in 3G networks using genetic algorithms," in *Proc. 3rd Int. Conf. Broadband Commun., Inf. Technol. Biomed. Appl.*, Nov. 2008, pp. 30–37.

[17] C. Edwards, "The future is femto—[comms femtocells]," *Eng. Technol.*, vol. 3, no. 15, pp. 70–73, Sep. 2008.

[18] *Broadband Evolved FEMTO Networks*, accessed on Nov. 17, 2016. [Online]. Available: http://www.ict-befemto.eu/

[19] L. C. Wang and C. J. Yeh, "3-cell network MIMO architectures with sectorization and fractional frequency reuse," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 6, pp. 1185–1199, Jun. 2011.

[20] *Energy Aware Radio and Network Technologies*, accessed on Nov. 16, 2016. [Online]. Available: https://www.ict-earth.eu/

[21] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, "A survey of self organisation in future cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 336–361, Feb. 2013.

[22] M. Jaber, Z. Dawy, N. Akl, and E. Yaacoub, "Tutorial on LTE/LTE–A cellular network dimensioning using iterative statistical analysis," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1355–1383, 2nd Quart., 2016.

[23] Z. Dawy, A. Husseini, E. Yaacoub, and L. Al-Kanj, "A wireless communications laboratory on cellular network planning," *IEEE Trans. Edu.*, vol. 53, no. 4, pp. 653–661, Nov. 2010.

[24] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. D. Fonseca, "Performance assessment of multiobjective optimizers: An analysis and review," *IEEE Trans. Evol. Comput.*, vol. 7, no. 2, pp. 117–132, Apr. 2003.

[25] Y. Wu and S. Pierre, "Base station positioning in third generation mobile networks," in *Proc. IEEE Can. Conf. Elect. Comput. Eng. (CCECE)*, vol. 1. May 2003, pp. 31–34.

[26] A. Imran, M. A. Imran, and R. Tafazolli, "A novel self organizing framework for adaptive frequency reuse and deployment in future cellular networks," in *Proc. IEEE 21st Int Pers. Indoor Mobile Radio Commun. Symp. (PIMRC)*, Sep. 2010, pp. 2354–2359.

[27] S. B. Jamaa, Z. Altman, J. M. Picard, and B. Fourestie, "Multi-objective strategies for automatic cell planning of UMTS networks," in *Proc. IEEE 59th Veh. Technol. Conf. (VTC)*, vol. 4. May 2004, pp. 2420–2424.

[28] C. Maple, L. Guo, and J. Zhang, "Parallel genetic algorithms for third generation mobile network planning," in *Proc. Int. Conf. Parallel Comput. Elect. Eng., (PARELEC)*, Sep. 2004, pp. 229–236.

[29] F. Gordejuela-Sanchez, A. Juttner, and J. Zhang, "A multiobjective optimization framework for IEEE 802.16e network design and performance analysis," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 2, pp. 202–216, Feb. 2009.

[30] A. F. Al Rawi, B. S. Sharif, and C. C. Tsimenidis, "Pareto-metaheuristic multi-objective network optimization for OFDMA-based systems," in *Proc. IEEE 6th Int. Conf. Wireless Mobile Comput., Netw. Commun. (WiMob)*, Oct. 2010, pp. 331–336.

[31] F. Gordejuela-Sanchez and J. Zhang, "LTE access network planning and optimization: A service-oriented and technology-specific perspective," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Nov. 2009, pp. 1–5.

[32] M. St-Hilaire, "Topological planning and design of UMTS mobile networks: A survey," *Wireless Commun. Mobile Comput.*, vol. 9, no. 7, pp. 948–958, Jul. 2009.

[33] L. Z. Ribeiro and L. A. DaSilva, "A framework for the dimensioning of broadband mobile networks supporting wireless Internet services," *IEEE Wireless Commun.*, vol. 9, no. 3, pp. 6–13, Jun. 2002.

[34] K. Tutschku and P. Tran-Gia, "Spatial traffic estimation and characterization for mobile communication network design," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 5, pp. 804–811, Jun. 1998.

[35] K. Tutschku, "Demand-based radio network planning of cellular mobile communication systems," in *Proc. 17th Annu. Joint Conf. IEEE Comput. Commun. Soc. (INFOCOM)*, vol. 3. Mar. 1998, pp. 1054–1061.

[36] R. Mathar and M. Schmeink, "Optimal base station positioning and channel assignment for 3G mobile networks by integer programming," *Ann. Oper. Res.*, vol. 107, nos. 1–4, pp. 225–236, Oct. 2001.

[37] R. M. Whitaker and S. Hurley, "Evolution of planning for wireless communication systems," in *Proc. 36th Annu. Hawaii Int. Conf. Syst. Sci.*, Jan. 2003, p. 10.

[38] R. Pattuelli and V. Zingarelli, "Precision of the estimation of area coverage by planning tools in cellular systems," *IEEE Pers. Commun.*, vol. 7, no. 3, pp. 50–53, Jun. 2000.

[39] C. Takahashi, Z. Yun, M. F. Iskander, G. Poilasne, V. Pathak, and J. Fabrega, "Propagation-prediction and site-planning software for wireless communication systems," *IEEE Antennas Propag. Mag.*, vol. 49, no. 2, pp. 52–60, Apr. 2007.

[40] U. Okumura, "Field strength and its variability in VHF and UHF land-mobile radio service," *Rev. Elect. Commun. Lab.*, vol. 16, nos. 9–10, pp. 825–873, 1968.

[41] M. Hatay, "Empirical formula for propagation loss in land mobile radio services," *IEEE Trans. Veh. Technol.*, vol. 29, no. 3, pp. 317–325, Aug. 1980.

[42] (1999). *COST 231 Final Report, Digital Mobile Radio Towards Future Generation Systems*. [Online]. Available: http:/www.lx.it.pt/cost231

[43] T. S. Rappaport *et al.*, "Millimeter wave mobile communications for 5G cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[44] R. R. Collmann, "Evaluation of methods for determining the mobile traffic distribution in cellular radio networks," *IEEE Trans. Veh. Technol.*, vol. 50, no. 6, pp. 1629–1635, Nov. 2001.

[45] M. Galota, C. Glaer, S. Reith, and H. Vollmer, "A polynomial-time approximation scheme for base station positioning in UMTS networks," in *Proc. 5th Int. Workshop Discrete Algorithms Methods Mobile Comput. Commun. (DIALM)*, New York, NY, USA, 2001, pp. 52–59.

[46] E. Amaldi, A. Capone, F. Malucelli, and F. Signori, "UMTS radio planning: Optimizing base station configuration," in *Proc. IEEE 56th Veh. Technol. Conf. (VTC)*, vol. 2. Sep. 2002, pp. 768–772.

[47] N. Weicker, G. Szabo, K. Weicker, and P. Widmayer, "Evolutionary multiobjective optimization for base station transmitter placement with frequency assignment," *IEEE Trans. Evol. Comput.*, vol. 7, no. 2, pp. 189–203, Apr. 2003.

[48] D. Tsilimantos, D. Kaklamani, and G. Tsoulos, "Particle swarm optimization for UMTS WCDMA network planning," in *Proc 3rd Int. Symp. Wireless Pervasive Comput. (ISWPC)*, May 2008, pp. 283–287.

[49] R. Mathar and T. Niessen, "Optimum positioning of base stations for cellular radio networks," *Wireless Netw.*, vol. 6, no. 6, pp. 421–428, Dec. 2000. [Online]. Available: http://dx.doi.org/10.1023/A:1019263308849

[50] A. Eisenblatter and H.-F. Geerd, "Wireless network design: Solution-oriented modeling and mathematical optimization," *IEEE Wireless Commun.*, vol. 13, no. 6, pp. 8–14, Dec. 2006.

[51] K. R. Guruprasad, "Generalized Voronoi partition: A new tool for optimal placement of base stations," in *Proc. IEEE 5th Int. Conf. Adv. Netw. Telecommun. Syst. (ANTS)*, Dec. 2011, pp. 1–3.

[52] A. Imran, E. Yaacoub, Z. Dawy, and A. Abu-Dayya, "A mathematical modeling approach and a novel solution for sector azimuth angle planning," in *Proc. IEEE Medit. Electrotech. Conf. (MELECON)*, Beirut, Lebanon, Apr. 2014, pp. 334–338.

[53] M. A. Aldajani, "Convolution-based placement of wireless base stations in urban environment," *IEEE Trans. Veh. Technol.*, vol. 57, no. 6, pp. 3843–3848, Nov. 2008.

[54] N. Lev-Tov and D. Peleg, "Exact algorithms and approximation schemes for base station placement problems," in *Proc. 8th Scandin. Workshop Algorithm Theory*, 2002, pp. 90–99. [Online]. Available: http://dl.acm.org/citation.cfm?id=645901.672620

[55] A. Imran, M. A. Imran, and R. Tafazolli, "Relay station access link spectral efficiency optimization through SO of Macro BS tilts," *IEEE Commun. Lett.*, vol. 15, no. 12, pp. 1326–1328, Dec. 2011.

[56] A. Imran, M. A. Imran, and R. Tafazolli, "Energy-aware adaptive sectorisation in LTE systems," in *Proc. IEEE 22nd Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2011, pp. 2402–2406.

[57] A. Imran, M. A. Imran, A.-U. Quddus, and R. Tafazolli, "Distributed spectral efficiency optimization at hotspots through self organisation of BS tilts," in *Proc. IEEE GLOBECOM Workshops*, Dec. 2011, pp. 570–574.

[58] C. Dimopoulos and A. M. S. Zalzala, "Recent developments in evolutionary computation for manufacturing optimization: Problems, solutions, and comparisons," *IEEE Trans. Evol. Comput.*, vol. 4, no. 2, pp. 93–113, Jul. 2000.

[59] E. Yaacoub and Z. Dawy, "LTE radio network planning with HetNets: BS placement optimization using simulated annealing," in *Proc. IEEE Medit. Electrotech. Conf. (MELECON)*, Beirut, Lebanon, Apr. 2014, pp. 327–333.

[60] H. P. Lin, R. T. Juang, D. B. Lin, C. Y. Ke, and Y. Wang, "Cell planning scheme for WCDMA systems using genetic algorithm and measured background noise floor," *IEE Proc.-Commun.*, vol. 151, no. 6, pp. 595–600, Dec. 2004.

[61] F. Gu, H. Liu, and M. Li, "Evolutionary algorithm for the radio planning and coverage optimization of 3G cellular networks," in *Proc. Int. Conf. Comput. Intell. Secur. (CIS)*, vol. 2. Dec. 2009, pp. 109–113.

[62] H. Yang, J. Wang, X. Song, Y. Yang, and M. Wang, "Wireless base stations planning based on GIS and genetic algorithms," in *Proc. 19th Int. Conf. Geoinformat.*, Jun. 2011, pp. 1–5.

[63] H. M. Elkamchouchi, H. M. Elragal, and M. A. Makar, "Cellular radio network planning using particle swarm optimization," in *Proc. Nat. Radio Sci. Conf. (NRSC)*, Mar. 2007, pp. 1–8.

[64] A. Awada, B. Wegmann, I. Viering, and A. Klein, "Optimizing the radio network parameters of the long term evolution system using Taguchi's method," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 3825–3839, Oct. 2011.

[65] V. Berrocal-Plaza, M. A. Vega-Rodríguez, J. A. Gómez-Pulido, and J. M. Sánchez-Pérez, "Artificial bee colony algorithm applied to WiMAX network planning problem," in *Proc. 11th Int. Conf. Intell. Syst. Design Appl. (ISDA)*, Nov. 2011, pp. 504–509.

[66] C. Y. Lee and H. G. Kang, "Cell planning with capacity expansion in mobile communications: A Tabu search approach," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1678–1691, Sep. 2000.

[67] I. Toros and P. Fazekas, "An energy efficient cellular mobile network planning algorithm," in *Proc. IEEE 73rd Veh. Technol. Conf. (VTC Spring)*, May 2011, pp. 1–5.

[68] B. Chamaret *et al.*, "Radio network optimization with maximum independent set search," in *Proc. IEEE 47th Veh. Technol. Conf.*, vol. 2. May 1997, pp. 770–774.

[69] Y. Wu and S. Pierre, "A new hybrid constraint-based approach for 3G network planning," *IEEE Commun. Lett.*, vol. 8, no. 5, pp. 277–279, May 2004.

[70] E. Amaldi, A. Capone, and F. Malucelli, "Optimizing base station siting in UMTS networks," in *Proc. IEEE 53rd Veh. Technol. Conf. (VTC Spring)*, vol. 4. May 2001, pp. 2828–2832.

[71] E. Amaldi, A. Capone, F. Malucelli, and F. Signori, "A mathematical programming approach for w-cdma radio planning with uplink and downlink constraints," in *Proc. IEEE 58th Veh. Technol. Conf. (VTC)*, vol. 2, Oct. 2003, pp. 806–810.

[72] E. Amaldi, A. Capone, and F. Malucelli, "Optimizing UMTS radio coverage via base station configuration," in *Proc. 13th IEEE Int. Symp. Pers., Indoor Mobile Radio Commun.*, vol. 1. Sep. 2002, pp. 315–319.

[73] E. Amaldi, A. Capone, F. Malucelli, and F. Signori, "Optimization models and algorithms for downlink UMTS radio planning," in *Proc. IEEE Wireless Commun. Netw. (WCNC)*, vol. 2. Mar. 2003, pp. 827–831.

[74] F. Athley, "On base station antenna beamwidth for sectorized WCDMA systems," in *Proc. IEEE 64th Veh. Technol. Conf. (VTC)*, Sep. 2006, pp. 1–5.

[75] S. Hurley, "Planning effective cellular mobile radio networks," *IEEE Trans. Veh. Technol.*, vol. 51, no. 2, pp. 243–253, Mar. 2002.

[76] S. U. Thiel, P. Giuliani, L. J. Ibbetson, and D. Lister, "An automated UMTS site selection tool," in *Proc. 3rd Int. Conf. 3G Mobile Commun. Technol.*, May 2002, pp. 69–73.

[77] D. Abusch-Magder, "Novel algorithms for reducing cell sites during a technology upgrade and network overlay," in *Proc. IEEE Wireless Commun. Netw. Conf.*, vol. 3. Mar. 2005, pp. 1726–1732.

[78] A. Molina, G. E. Athanasiadou, and A. R. Nix, "The automatic location of base-stations for optimised cellular coverage: A new combinatorial approach," in *Proc. IEEE 49th Veh. Technol. Conf.*, vol. 1. Jul. 1999, pp. 606–610.

[79] J. Zhang, L. Guo, and J. Y. Wu, "An integrated approach for UTRAN planning and optimization," in *Proc. IEEE 59th Veh. Technol. Conf. (VTC)*, vol. 4. May 2004, pp. 2360–2364.

[80] S. Sohn and G.-S. Jo, "Optimization of base stations positioning in mobile networks," in *Proc. Int. Conf. Comput. Sci. Appl. (ICCSA)*, 2006, pp. 779–787. [Online]. Available: http://dx.doi.org/10.1007/11751588-81

[81] Y. H. Chew, A. C. W. Tan, B. S. Yeo, and C. Wang, "A cell planning approach in non-isotropic transmission medium with the capability of supporting cellular networks downsizing," in *Proc. IEEE 64th Veh. Technol. Conf. (VTC)*, Sep. 2006, pp. 1–5.

[82] T. Tsourakis and K. Voudouris, "WiMAX network planning and system's performance evaluation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2007, pp. 1948–1953.

[83] S. Priem-Mendes, J. A. Gomez-Pulidom, M. A. Vega-Rodriguez, A. M. Pereira, and J. M. Sanchez Perez, "Fast wide area network design optimisation using differential evolution," in *Proc. Int. Conf. Adv. Eng. Comput. Appl. Sci. (ADVCOMP)*, Nov. 2007, pp. 3–10.

[84] Y. Yu, S. Murphy, and L. Murphy, "A clustering approach to planning base station and relay station locations in IEEE 802.16j multi-hop relay networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2008, pp. 2586–2591.

[85] L. Zeng, L. Wang, and C. Ding, "Site selection for wireless base station based on map partitioning," in *Proc. 4th Int. Conf. Wireless Commun., Netw. Mobile Comput. (WiCOM)*, Oct. 2008, pp. 1–4.

[86] A. Kamar, S. J. Nawaz, M. Patwary, M. Abdel-Maguid, and S.-U.-R. Qureshi, "Optimized algorithm for cellular network planning based on terrain and demand analysis," in *Proc. 2nd Int. Conf. Comput. Technol. Develop. (ICCTD)*, Nov. 2010, pp. 359–364.

[87] L. Hu, I. Z. Kovacs, P. Mogensen, O. Klein, and W. Stormer, "Optimal new site deployment algorithm for heterogeneous cellular networks," in *Proc. IEEE Veh. Technol. Conf. (VTC Fall)*, Sep. 2011, pp. 1–5.

[88] A. A. Khalek, L. Al-Kanj, Z. Dawy, and G. Turkiyyah, "Site placement and site selection algorithms for UMTS radio planning with quality constraints," in *Proc. IEEE 17th Int. Conf. Telecommun. (ICT)*, Apr. 2010, pp. 375–381.

[89] A. A. Khalek, L. Al-Kanj, Z. Dawy, and G. Turkiyyah, "Optimization models and algorithms for joint uplink/downlink UMTS radio network planning with SIR-based power control," *IEEE Trans. Veh. Technol.*, vol. 60, no. 4, pp. 1612–1625, May 2011.

[90] E. Oh and B. Krishnamachari, "Energy savings through dynamic base station switching in cellular wireless access networks," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2010, pp. 1–5.

[91] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, "Toward dynamic energy-efficient operation of cellular network infrastructure," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 56–61, Jun. 2011.

[92] P. Gonzalez-Brevis *et al.*, "Base station location optimization for minimal energy consumption in wireless networks," in *Proc. IEEE 73rd Veh. Technol. Conf. (VTC Spring)*, May 2011, pp. 1–5.

[93] B. Rengarajan, G. Rizzo, and M. Marsan, "Bounds on QoS-constrained energy savings in cellular access networks with sleep mode," in *Proc. 23rd Int. Teletraffic Congr. (ITC)*, Sep. 2011, pp. 47–54.

[94] M. A. Marsan, L. Chiaraviglio, D. Ciullo, and M. Meo, "Optimal energy savings in cellular access networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2009, pp. 1–5.
[95] Z. Niu, Y. Wu, J. Gong, and Z. Yang, "Cell zooming for cost-efficient green cellular networks," *IEEE Commun. Mag.*, vol. 48, no. 11, pp. 74–79, Nov. 2010.
[96] F. Richter, G. Fettweis, M. Gruber, and O. Blume, "Micro base stations in load constrained cellular mobile radio networks," in *Proc. IEEE 21st Int. Symp. Pers., Indoor Mobile Radio Commun. Workshops (PIMRC Workshops)*, Sep. 2010, pp. 357–362.
[97] W. El-Beaino, A. M. El-Hajj, and Z. Dawy, "A proactive approach for LTE radio network planning with green considerations," in *Proc. 19th Int. Conf. Telecommun. (ICT)*, Apr. 2012, pp. 1–5.
[98] M. Katsigiannis and H. Hämmäinen, "Energy and cost efficient radio access network deployments—Case Finland," in *Proc. 4th Joint IFIP Wireless Mobile Netw. Conf. (WMNC)*, Oct. 2011, pp. 1–8.
[99] Y. Qi, M. Imran, and R. Tafazolli, "On the energy aware deployment strategy in cellular systems," in *Proc. IEEE 21st Int. Symp. Pers., Indoor Mobile Radio Commun. Workshops (PIMRC Workshops)*, Sep. 2010, pp. 363–367.
[100] M. A. Marsan and M. Meo, "Energy efficient management of two cellular access networks," *SIGMETRICS Perform. Eval. Rev.*, vol. 37, no. 4, pp. 69–73, Mar. 2010. [Online]. Available: http://doi.acm.org/10.1145/1773394.1773406
[101] S. Boiardi, A. Capone, and B. Sansó, "Planning for energy-aware wireless networks," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 156–162, Feb. 2014.
[102] A. M. C. A. Koster, M. Kutschka, and C. Raack, "Towards robust network design using integer linear programming techniques," in *Proc. 6th EURO-NF Conf. Next Generat. Internet (NGI)*, Jun. 2010, pp. 1–8.
[103] Y. H. Chew, R. Mo, and B. S. Yeo, "Inclusion of mobility in cell planning: Multi-period joint optimization," in *Proc. IEEE 70th Veh. Technol. Conf. Fall (VTC Fall)*, Sep. 2009, pp. 1–4.
[104] S. Yang and F. A. Kuipers, "Traffic uncertainty models in network planning," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 172–177, Feb. 2014.
[105] S. E. Terblanche and R. Wessäly, and J. M. Hattingh, "Lagrangian relaxation as a solution approach to solving the survivable multi-hour network design problem," in *Proc. South African Telecommun. Netw. Appl. Conf. (SATNAC)*, 2007, pp. 1–5.
[106] G. Oriolo, "Domination between traffic matrices," *Math. Oper. Res.*, vol. 33, no. 1, pp. 91–96, 2008.
[107] W. Ben-Ameur and H. Kerivin, "New economical virtual private networks," *Commun. ACM*, vol. 46, no. 6, pp. 69–73, Jun. 2003.
[108] N. G. Duffield, P. Goyal, A. G. Greenberg, P. P. Mishra, K. K. Ramakrishnan, and J. E. van der Merive, "A flexible model for resource management in virtual private networks," in *Proc. ACM SIGCOMM*, 1999, pp. 95–108.
[109] A. L. Soyster, "Technical Note—Convex programming with set-inclusive constraints and applications to inexact linear programming," *Oper. Res.*, vol. 21, no. 5, pp. 1154–1157, 1973.
[110] T. Bauschert, C. Büsing, F. D'Andreagiovanni, A. C. A. Koster, M. Kutschka, and U. Steglich, "Network planning under demand uncertainty with robust optimization," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 178–185, Feb. 2014.
[111] G. Claßen, A. M. C. A. Koster, and A. Schmeink, "Robust planning of green wireless networks," in *Proc. 5th Int. Conf. Netw. Games, Control Optim. (NetGCooP)*, Oct. 2011, pp. 1–5.
[112] I. Garcia, N. Kusashima, K. Sakaguchi, K. Araki, S. Kaneko, and Y. Kishi, "Impact of base station cooperation on cell planning," *EURASIP J. Wireless Commun. Netw.*, vol. 2010, Dec. 2010, Art. no. 406749. [Online]. Available: http://jwcn.eurasipjournals.com/content/2010/1/406749
[113] D. Cao, S. Zhou, C. Zhang, and Z. Niu, "Energy saving performance comparison of coordinated multi-point transmission and wireless relaying," in *Proc. IEEE Global Telecommun. Conf. (GLOBECOM)*, Dec. 2010, pp. 1–5.
[114] Z. Niu, S. Zhou, Y. Hua, Q. Zhang, and D. Cao, "Energy-aware network planning for wireless cellular system with inter-cell cooperation," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1412–1423, Apr. 2012.
[115] Q. T. Zhang, "Bridging the gap between dynamic and static methods for cell planning," *IEEE Trans. Veh. Technol.*, vol. 50, no. 5, pp. 1224–1230, Sep. 2001.
[116] R. Giuliano and F. Mazzenga, "Exponential effective SINR approximations for OFDM/OFDMA-based cellular system planning," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4434–4439, Sep. 2009.
[117] K. Majewski, U. Turke, X. Huang, and B. Bonk, "Analytical cell load assessment in OFDM radio networks," in *Proc. IEEE 18th Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2007, pp. 1–5.
[118] I. Siomina and D. Yuan, "Analysis of cell load coupling for LTE network planning and optimization," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2287–2297, Jun. 2012.
[119] I. Siomina, A. Furuskär, and G. Fodor, "A mathematical framework for statistical QoS and capacity studies in OFDM networks," in *Proc. IEEE 20th Int. Symp. Pers., Indoor Mobile Radio Commun.*, Sep. 2009, pp. 2772–2776.
[120] F. J. Velez, M. del Camino Noguera, O. Holland, and H. Aghvami, "Fixed WiMAX profit maximisation with energy saving through relay sleep modes and cell zooming," *J. Green Eng.*, vol. 1, no. 4, pp. 355–381, Jun. 2011.
[121] H. Tabassum, F. Yilmaz, Z. Dawy, and M.-S. Alouini, "A framework for uplink intercell interference modeling with channel-based scheduling," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 206–217, Jan. 2013.
[122] H. Tabassum, F. Yilmaz, Z. Dawy, and M.-S. Alouini, "A statistical model of uplink inter-cell interference with slow and fast power control mechanisms," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3953–3966, Sep. 2013.
[123] R. K. Taplin, D. M. Ryan, S. M. Allen, S. Hurley, and N. J. Thomas, "Algorithms for the automatic design of WiMAX networks," in *Proc. Int. Conf. Next Generat. Mobile Appl., Services Technol. (NGMAST)*, Sep. 2007, pp. 322–327.
[124] M. Werner *et al.*, "Cost assessment and optimization methods for multi-node radio access networks," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, May 2008, pp. 2601–2605.
[125] G. Micallef, P. Mogensen, H.-O. Scheck, and E. Lang, "Energy efficient evolution of mobile networks: Macro-only upgrades vs. a joint-pico deployment strategy," in *Proc. IEEE 73rd Veh. Technol. Conf. (VTC Spring)*, May 2011, pp. 1–5.
[126] W. El-Beaino, A. M. El-Hajj, and Z. Dawy, "A proactive approach for lte radio network planning with green considerations," in *Proc. 19th Int. Conf. Telecommun. (ICT)*, Apr. 2012, pp. 1–5.
[127] J. G. Andrews, "Seven ways that HetNets are a cellular paradigm shift," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 136–144, Mar. 2013.
[128] X. Ge, H. Cheng, M. Guizani, and T. Han, "5G wireless backhaul networks: Challenges and research advances," *IEEE Netw.*, vol. 28, no. 6, pp. 6–11, Nov. 2014.
[129] N. Bhushan *et al.*, "Network densification: The dominant theme for wireless evolution into 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 82–89, Feb. 2014.
[130] T. Han and N. Ansari, "On greening cellular networks via multicell cooperation," *IEEE Wireless Commun.*, vol. 20, no. 1, pp. 82–89, Feb. 2013.
[131] S. Navaratnarajah, A. Saeed, M. Dianati, and M. A. Imran, "Energy efficiency in heterogeneous wireless access networks," *IEEE Wireless Commun.*, vol. 20, no. 5, pp. 37–43, Oct. 2013.
[132] G. Piro *et al.*, "Hetnets powered by renewable energy sources: Sustainable next-generation cellular networks," *IEEE Internet Comput.*, vol. 17, no. 1, pp. 32–39, Jan. 2013.
[133] O. Onireti, A. Imran, M. A. Imran, and R. Tafazolli, "Energy efficient inter-frequency small cell discovery in heterogeneous networks," *IEEE Trans. Veh. Technol.*, vol. 65, no. 9, pp. 7122–7135, Sep. 2015.
[134] I. Siomina and D. Yuan, "Optimization Approaches for Planning Small Cell Locations in Load-Coupled Heterogeneous LTE Networks," in *Proc. IEEE Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2013, pp. 2904–2908.
[135] A. B. Saleh, Bulakci, J. Hämäläinen, S. Redana, and B. Raaf, "Analysis of the impact of site planning on the performance of relay deployments," *IEEE Trans. Veh. Technol.*, vol. 61, no. 7, pp. 3139–3150, Sep. 2012.
[136] O. Bulakci, S. Redana, B. Raaf, and J. Hämäläinen, "Performance enhancement in LTE-Advanced relay networks via relay site planning," in *Proc. IEEE 71st Veh. Technol. Conf. (VTC Spring)*, May 2010, pp. 1–5.
[137] W. Guo and T. O'Farrell, "Relay deployment in cellular networks: Planning and optimization," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 8, pp. 1597–1606, Aug. 2013.

[138] J. Hoydis and M. Debbah, "Green, cost-effective, flexible, small cell networks," *IEEE Commun. Soc.*, vol. 5, no. 5, pp. 23–26, Sep. 2010.
[139] A. Maeder *et al.*, "Towards a flexible functional split for cloud-RAN networks," in *Proc. Eur. Conf. Netw. Commun. (EuCNC)*, Jun. 2014, pp. 1–5.
[140] A. Zakrzewska *et al.*, "A framework for joint optical-wireless resource management in multi-RAT, heterogeneous mobile networks," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC)*, Jun. 2013, pp. 895–899.
[141] B. Lin, X. Pan, R. He, and S. Li, "Joint wireless-optical infrastructure deployment and layout planning for cloud-radio access networks," in *Proc. Int. Wireless Commun. Mobile Comput. Conf. (IWCMC)*, Aug. 2014, pp. 1027–1032.
[142] Y. He, E. Dutkiewicz, G. Fang, and J. Shi, "Downlink capacity in cloud radio access networks with fractional frequency reuse," in *Proc. Int. Symp. Wireless Pers. Multimedia Commun. (WPMC)*, Sep. 2014, pp. 424–428.
[143] Y. Zhang and Y. J. Zhang, "User-centric virtual cell design for Cloud Radio Access Networks," in *Proc. IEEE 15th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2014, pp. 249–253.
[144] S.-Y. Lien, K.-C. Chen, and Y. Lin, "Toward ubiquitous massive accesses in 3GPP machine-to-machine communications," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 66–74, Apr. 2011.
[145] T. Taleb and A. Kunz, "Machine type communications in 3GPP networks: Potential, challenges, and solutions," *IEEE Commun. Mag.*, vol. 50, no. 3, pp. 178–184, Mar. 2012.
[146] A. Ksentini, Y. Hadjadj-Aoul, and T. Taleb, "Cellular-based machine-to-machine: Overload control," *IEEE Netw.*, vol. 26, no. 6, pp. 54–60, Jan. 2013.
[147] "Service requirements for machine type communication (MTC)," 3rd Generation Partnership Project (3GPP), Tech. Rep. TS 22.368.
[148] "Study on provision of low-cost machine type communications (MTC) user equipments (UEs) based on LTE," 3rd Generation Partnership Project (3GPP), Tech. Rep. TR 36.888 V12.0.0, 2013.
[149] M. Hasan, E. Hossain, and D. Niyato, "Random access for machine-to-machine communication in LTE-advanced networks: Issues and approaches," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 86–93, Jun. 2013.
[150] K. Zheng, F. Hu, W. Wang, W. Xiang, and M. Dohler, "Radio resource allocation in LTE-advanced cellular networks with M2M communications," *IEEE Commun. Mag.*, vol. 50, no. 7, pp. 184–192, Jul. 2012.
[151] C. Y. Ho and C.-Y. Huang, "Energy-saving massive access control and resource allocation schemes for M2M communications in OFDMA cellular networks," *IEEE Wireless Commun. Lett.*, vol. 1, no. 3, pp. 209–212, Jun. 2012.
[152] R. Schneiderman, "Machine-to-machine connectivity market is booming [special reports]," *IEEE Signal Process. Mag.*, vol. 30, no. 4, pp. 10–13, Jul. 2013.
[153] D. Katusic, M. Weber, I. Bojic, G. Jezic, and M. Kusek, "Market, standardization, and regulation development in machine-to-machine communications," in *Proc. 12th Int. Conf. Softw., Telecommun., Comput. Netw. (SoftCOM)*, 2012, pp. 1–7.
[154] C. Ma, J. Yue, H. Yu, H. Luo, W. Zhou, and X. Sun, "An interference coordination mechanism based on resource allocation for network controlled device-to-device communication," in *Proc. IEEE/CIC ICCC*, Xi'an, China, Aug. 2013, pp. 109–114.
[155] F. Wang, L. Song, Z. Han, Q. Zhao, and X. Wang, "Joint scheduling and resource allocation for device-to-device underlay communication," in *Proc. IEEE WCNC*, Shanghai, China, Apr. 2013, pp. 134–139.
[156] H. Wang and X. Chu, "Distance-constrained resource-sharing criteria for device-to-device communications underlaying cellular networks," *Electron. Lett.*, vol. 48, no. 9, pp. 528–530, Apr. 2012.
[157] D. H. Lee, K. W. Choi, W. S. Jeon, and D. G. Jeong, "Resource allocation scheme for device-to-device communication for maximizing spatial reuse," in *Proc. IEEE WCNC*, Shanghai, China, Apr. 2013, pp. 112–117.
[158] A. Mohamed, O. Onireti, M. A. Imran, A. Imran, and R. Tafazolli, "Control-data separation architecture for cellular radio access networks: A survey and outlook," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 446–465, 1st Quart., 2015.
[159] A. Capone, A. F. dos Santos, I. Filippini, and B. Gloss, "Looking beyond green cellular networks," in *Proc. 9th Annu. Conf. Wireless Demand Netw. Syst. Services (WONS)*, Jan. 2012, pp. 127–130.
[160] H. Ishii, Y. Kishiyama, and H. Takahashi, "A novel architecture for LTE-B :C-plane/U-plane split and phantom cell concept," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2012, pp. 624–630.
[161] X. Xu, G. He, S. Zhang, Y. Chen, and S. Xu, "On functionality separation for green mobile networks: Concept study over LTE," *IEEE Commun. Mag.*, vol. 51, no. 5, pp. 82–90, May 2013.

AZAR TAUFIQUE received the B.Sc. degree (Hons.) in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2007, and the master's degree in telecommunications engineering from the University of Texas at Dallas, in 2010. He is currently pursuing the Ph.D. degree with the BSON Laboratory, The University of Oklahoma, Tulsa, OK, USA. He is a Telecommunications Technical Trainer, teaching 4G LTE courses to over 2500 Professionals at Verizon, T-Mobile, Sprint, and ATT, North America. His telecommunications blog is read by hundreds of ICT Professionals weekly. His research interests are in mobility management and new architectures for mobility management support in 5G.

MONA JABER received the B.E. degree in computer and communications engineering and the M.E. degree in electrical and computer engineering from the American University of Beirut, Beirut, Lebanon, in 1996 and 2014, respectively. She is currently pursuing the Ph.D. degree with the 5G Innovation Center, University of Surrey, U.K., researching on 5G backhaul solutions. She was a Telecommunication Consultant in various international firms with a focus on radio design of cellular networks, including GSM, GPRS, UMTS, and HSPA. Her research interests are wireless communications, cellular technologies, backhaul and fronthaul solutions, and self-optimization techniques.

ALI IMRAN is currently an Assistant Professor in telecommunications and the Director of the BSON Laboratory, at The University of Oklahoma, Tulsa, OK, USA. He led a multinational 1.045 million USD Research Project on Self Organizing Cellular Networks, QSON. He is also leading two NSF funded Research Projects on Enabling Ultra-Dense Future Cellular Networks, and Agile and Scalable Self-Healing Functionalities for Ultra Dense Future Cellular Networks. His current research interests include self-organizing networks, radio resource management, and big data analytics. He has authored over 40 peer-reviewed articles and has presented number of tutorials at international forums, such as the IEEE ICC, the IEEE WCNC, the European Wireless, and the CrownCom, on these topics. He is an Associate Fellow of the Higher Education Academy, U.K., and a member of Advisory Board to Special Technical Community on Big Data with the IEEE Computer Society.

ZAHER DAWY received the B.E. degree in computer and communications engineering from the American University of Beirut (AUB), Beirut, Lebanon, in 1998, and the M.E. and Dr.-Ing. degrees in communications engineering from the Munich University of Technology, Munich, Germany, in 2000 and 2004, respectively. Since 2004, he was with the Department of Electrical and Computer Engineering, AUB, where he is currently a Professor. His research and teaching interests include wireless communications, cellular technologies, context-aware mobile computing, mobile solutions for smart cities, computational biology, and biomedical engineering. He received the Abdul Hameed Shoman Award for Young Arab Researchers, in 2012, the IEEE Communications Society 2011 Outstanding Young Researcher Award in Europe, Middle East, and Africa Region, the AUB Teaching Excellence Award in 2008, the Best Graduate Award from TUM in 2000, the Youth and Knowledge Siemens Scholarship for Distinguished Students, in 1999, and the Distinguished Graduate Medal of Excellence from Hariri Foundation, in 1998. He was an Executive Editor of the *Wiley Transactions on Emerging Telecommunications Technologies* from 2011 to 2014. He is an Editor of the IEEE Communications Surveys and Tutorials, the IEEE Transactions on Communications, the IEEE Transactions on Wireless Communications, and the *Elsevier Physical Communications*.

ELIAS YAACOUB received the B.E. degree in electrical engineering from Lebanese University, in 2002, and the M.E. degree in computer and communications engineering and the Ph.D. degree in electrical and computer engineering from the American University of Beirut (AUB), in 2005 and 2010, respectively. He was a Research Assistant with AUB from 2004 to 2005 and the Munich University of Technology in 2005. From 2005 to 2007, he was a Telecommunications Engineer with Dar Al-Handasah, Shair and Partners. In 2014, he was a Research Scientist/ Research and Development Expert with the Qatar Mobility Innovations Center. He joined the Strategic Decisions Group, where he was a Consultant until 2016. He is currently an Associate Professor with the Arab Open University. His current research interests include wireless communications, resource allocation in wireless networks, intercell interference mitigation techniques, antenna theory, sensor networks, and physical layer security.